(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 May 2002 (10.05.2002)

PCT

(10) International Publication Number
WO 02/36784 A1

(51) International Patent Classification[7]: C12N 15/63, 5/10, C12Q 1/68, A01K 67/027, A61K 49/00

(21) International Application Number: PCT/AU01/01407

(22) International Filing Date: 1 November 2001 (01.11.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

PR 1161	1 November 2000 (01.11.2000)	AU
PR 4901	10 May 2001 (10.05.2001)	AU

(71) Applicant (*for all designated States except US*): THE UNIVERSITY OF SYDNEY [AU/AU]; Parramatta Road, Sydney, New South Wales 2006 (AU).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): **LIDDLE, Christopher** [AU/AU]; 7 Hillside Street, Chatswood, New South Wales 2067 (AU). **GOODWIN, Bryan, James** [AU/US]; Apartment B10, Sterling Bluff, 140 BPW Club Road, Carrboro, NC 27510 (US). **ROBERTSON, Graham** [AU/AU]; c/o The University of Sydney, Parramatta Road, Sydney, New South Wales 2006 (AU).

(74) Agent: GRIFFITH HACK; GPO Box 4164, SYDNEY, New South Wales 2001 (AU).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/36784 A1

(54) Title: TRANSGENIC ANIMALS FOR ANALYSING CYP3A4 CYTOCHROME P450 GENE REGULATION

(57) Abstract: The invention relates to the generation of non-human transgenic animals comprising a reporter construct for producing a detectable amount of a reporter molecule operably linked to a transcriptional regulatory nucleic acid molecule from the human CYP3A4 gene located between the initiation of transcription site of the gene and a position located 13,000 nucleotides upstream from the site. The invention also relates to the use of these animals for determining the effect of a compound, particularly, but not exclusively, a xenobiotic or steriod, on the regulation of expression of the CYP3A4 gene in a human.

TRANSGENIC ANIMALS FOR ANALYSING CYP3A4 CYTOCHROME P450 GENE REGULATION

TECHNICAL FIELD OF THE INVENTION

The invention relates to the generation of a transgenic animal and to the use of the animal for determining the effect of a compound, particularly, but not exclusively, a xenobiotic or steroid, on the regulation of expression of a P450 gene in a human.

BACKGROUND OF THE INVENTION

Many endogenous and exogenous compounds are observed to have a therapeutic effect in drug development trials *in vitro*. However, the intended therapeutic effect is often not realised in clinical practice, for example, when compounds are co-administered, because certain compounds induce the expression of the CYP3A4 gene. This induction generates CYP3A4 cytochrome P450 molecules which metabolise compounds before the intended therapeutic effect of each compound can be realised. Accordingly, induction of expression of the CYP3A4 gene interferes with intended dosage, leading to therapeutic failure or sub-optimal treatment.

Induction of CYP3A4 gene expression is a significant problem for drug development because time, resources and expense are wasted in the development of candidate drugs for therapy of particular disease conditions which will ultimately fail or perform sub-optimally in clinical practice.

It would be advantageous to have an animal model for use in drug development trials from which, at an early stage of drug development, one could determine whether a

candidate drug would be likely to achieve an intended therapeutic effect in a human.

Such an animal model would not be useful unless at least some of the aspects of the regulation of CYP3A4 gene expression in the human, especially tissue specific expression, are reproduced. This is because in the human, the CYP3A4 gene is expressed in specific tissues, including liver and small intestine, which many compounds inevitably come into contact with when administered for the purpose of therapy. Accordingly, one would be unable to determine whether the bio-availability of a candidate drug would be sufficient for achieving an intended therapeutic effect in clinical practice in a model which does not reproduce the constitutive and xenobiotic induced tissue specific expression of the CYP3A4 gene that is observed in the human.

WO99/61622 and Goodwin et al. 1999 disclose a nucleic acid molecule located 8 kb upstream from the initiation of transcription site of the CYP3A4 gene which regulates transcription of the CYP3A4 gene in response to xenobiotic compounds. These documents do not disclose elements for regulating the constitutive and xenobiotic inducible tissue specific and developmental expression of the CYP3A4 gene observed in a human.

There is a need for an animal model which reproduces at least some aspects of the expression of the CYP3A4 gene in a human, for determining whether a compound, for example, one identified in a drug development trial, would be likely to induce CYP3A4, and hence cause drug-drug interactions, or auto-induction of the metabolism of the drug under study.

DESCRIPTION OF THE INVENTION

The invention seeks to address the above identified need and in a first aspect provides a non-human mammal comprising:

(a) a regulatory nucleic acid molecule which is capable of regulating transcription of the human CYP3A4 gene and which comprises a nucleotide sequence that is identical to a sequence of the human CYP3A4 gene located between the initiation of transcription site of the gene and a position located at least 13,000 nucleotides upstream from the site; and

(b) a reporter nucleic acid molecule for producing a detectable amount of a reporter molecule for indicating regulation of transcription of the reporter nucleic acid molecule by the regulatory nucleic acid molecule

wherein the reporter and regulatory nucleic acid molecules are arranged to permit the regulatory nucleic acid molecule to regulate transcription of the reporter nucleic acid molecule.

As described herein, the inventors have found that the incorporation of a region of the human CYP3A4 gene that is located between the initiation of transcription site of the gene and a position 13,000 nucleotides upstream of the initiation of transcription site into an animal model provides the animal with sufficient genetic information for reproducing the constitutive and xenobiotic induced tissue specific expression of the CYP3A4 gene that is observed in humans. More specifically, the inventors have generated animal models which contain a transgene comprising this region and have observed that these models provide constitutive and xenobiotic inducible expression of a transgene in a tissue pattern which reproduces the

tissue specific expression of CYP3A4 which is observed in a human. Importantly, the level of constitutive expression is sufficient to allow one to observe the effect on the regulation of tissue specific transgene expression, of administration of a compound, for example, a xenobiotic or steroid, to the animal.

Further, the inventors have observed that the animal models described herein also reproduce aspects of the constitutive and xenobiotic inducible developmental expression of the CYP3A4 gene that is observed in humans.

These findings are unanticipated because prior to the invention, there was no suggestion that the genetic information required for simulating the constitutive and xenobiotic induced tissue specific or developmental expression of the CYP3A4 gene that is observed in a human would be contained in the region of the human CYP3A4 gene between the initiation of transcription site of the gene and a position 13,000 nucleotides upstream of the initiation of transcription site.

Further, prior to the invention, differences in the induction profile of the mouse CYP3A11 and the human CYP3A4 gene had been observed, and differences had also been observed in the ligand binding profile of mouse transcription factors, especially PXR and CAR, and human PXR and CAR. Accordingly, there was no suggestion that a non-human animal would have factors sufficient for interacting with a region of the CYP3A4 gene for reproducing the constitutive and xenobiotic induced tissue specific or developmental expression of CYP3A4 observed in a human.

Further, prior to the invention, mechanisms associated with transgene integration had been observed, such as gene

silencing and mosaic transgene expression which limited the extent to which an a transcriptional enhancer element incorporated into a trangenic model could reproduce regulation of gene expression observed in a human. Accordingly, there was no suggestion that a region of the human CYP3A4 gene would be capable of reproducing the regulation of expression of the CYP3A4 gene that is observed in a human. However, as described herein, the inventors have shown in 2 separate founder lines that the expression of the transgene reproduces aspects of CYP3A4 gene expression that are observed in humans.

Thus in a second aspect, the invention provides a non human mammal comprising:

(a) a regulatory nucleic acid molecule comprising a nucleotide sequence that is identical to the nucleotide sequence of the human CYP3A4 gene that extends about 13,000 nucleotides upstream from the initiation of transcription site of the gene; and

(b) a reporter nucleic acid molecule for producing a detectable amount of a reporter molecule for indicating regulation of transcription of the reporter nucleic acid molecule by the regulatory nucleic acid molecule

wherein the reporter and regulatory nucleic acid molecules are arranged to permit the regulatory nucleic acid molecule to regulate transcription of the reporter nucleic acid molecule.

In one embodiment, the regulatory nucleic acid molecule comprises the sequence shown in SEQ ID NO:1.

Further, as described herein, the inventors have generated transgenic animals which contain a region of the human

CYP3A4 gene between the initiation of transcription site and a position about 3,200 nucleotides upstream of the initiation transcription site and observed that the transgene is not constitutively expressed or inducible by xenobiotics in these animals. Accordingly, the inventors have found that the genetic information required for reproducing the constitutive and xenobiotic induced tissue specific and developmental expression of CYP3A4 observed in a human is contained in the region of the human CYP3A4 gene between the position located about 3,200 nucleotides upstream of the initiation of transcription site of the gene and a position 13,000 nucleotides upstream of the initiation of transcription site.

Thus, in a third aspect, the invention provides a non-human mammal comprising:

(a) a regulatory nucleic acid molecule comprising a nucleotide sequence that is identical to the sequence of the human CYP3A4 gene that extends about 8,000 nucleotides upstream from a position about 3,000 nucleotides upstream from the initiation of transcription site of the gene; and

(b) a reporter nucleic acid molecule for producing a detectable amount of a reporter molecule for indicating regulation of transcription of the reporter nucleic acid molecule by the regulatory nucleic acid molecule

wherein the reporter and regulatory nucleic acid molecules are arranged to permit the regulatory nucleic acid molecule to regulate transcription of the reporter nucleic acid molecule.

In one embodiment, the regulatory nucleic acid molecule comprises the sequence shown in SEQ ID NO:2.

In a fourth aspect, the invention provides a non-human mammal comprising:

(a) a regulatory nucleic acid molecule which is capable of regulating transcription of the human CYP3A4 gene and which comprises a nucleotide sequence that is identical to the sequence of the human CYP3A4 gene that extends about 600 nucleotides upstream from a position about 7,200 nucleotides upstream of the initiation of transcription site of the gene; and

(b) a reporter nucleic acid molecule for producing a detectable amount of a reporter molecule for indicating regulation of transcription of the reporter nucleic acid molecule by the regulatory nucleic acid molecule

wherein the reporter and regulatory nucleic acid molecules are arranged to permit the regulatory nucleic acid molecule to regulate transcription of the reporter nucleic acid molecule.

In one embodiment, the regulatory nucleic acid molecule comprises the sequence shown in SEQ ID NO:3.

In another embodiment, the regulatory nucleic acid molecule has the sequence of any one of the following fragments of the CYP3A4 gene:

(i) a fragment consisting of from nucleotide positions -13,000 to +53;

(ii) a fragment consisting of from nucleotide positions -13,000 to -12,700 contiguous with -8000 to +53;

(iii) a fragment consisting of from nucleotide positions -13,000 to -5,100 contiguous with -1,200 to +53;

(v) a fragment consisting of from nucleotide positions -7,800 to -6,000 contiguous with -362 to +53;

(vi) a fragment consisting of from nucleotide positions

-7,500 to -6,000 contiguous with -362 to +53;

A regulatory nucleic acid molecule which has the sequence of a fragment consisting of from nucleotide positions -7836 to -7207 contiguous with -362 to +53 is particularly preferred, as this construct contains the minimal sequences necessary for regulating transcription of the human CYP3A4 gene, more specifically, an element responsive to xenobiotics (the "Xenobiotic Response Element Module" or "XREM") and the proximal promoter of the CYP3A4 gene.

The regulatory nucleic acid molecule of the invention typically contains at least one enhancer capable of regulating transcription of a human CYP3A4 gene when contacted with a nuclear receptor. Examples of such enhancers are those capable of regulating transcription of a human CYP3A4 gene when contacted with a nuclear receptor bound to a ligand, such as a xenobiotic or steroid. Other examples are those capable of regulating transcription of a human CYP3A4 gene when contacted with a nuclear receptor consisting of a heterodimer of PXR (pregnane X receptor, otherwise known as SXR (steroid and xenobiotic receptor)) and RXR (9-cis retinoic acid receptor), or CAR (constitutive androstane receptor-β) and RXR.

The inventors believe that certain nucleic acid molecules which have substantially the same nucleotide sequence as a regulatory nucleic acid molecule of the invention would also have sufficient genetic information for reproducing the constitutive and xenobiotic induced tissue specific and developmental expression of the CYP3A4 gene that is observed in a human. Accordingly, it will be understood that nucleotides could be modified or deleted in regions of the regulatory nucleic acid molecule, more specifically, those regions which do not contain an

enhancer such as those described above, without significantly limiting the capacity of the molecule to regulate transcription of the human CYP3A4 gene.

The inventors recognise that it would be advantageous to provide an animal model further capable of reproducing the expression of other human genes, specifically those genes encoding products which modify or modulate the therapeutic activity of exogenous and endogenous compounds used for therapy and cause drug-drug interactions, for example, cytochrome P450 genes or ABC transporter superfamily genes, for example, P-glycoprotein (otherwise known as MDR-1). The regions controlling the constitutive and xenobiotic induced tissue specific expression of some of these genes are known, and in some instances, non-human animal models have been generated. The inventors recognise that the genetic background of these animals could be incorporated into the non-human mammal of the present invention, for example, by conventional breeding techniques.

Thus in a fifth aspect, the invention provides a non-human mammal of any one of the first to fourth aspects of the invention, further comprising:

(c) a further regulatory nucleic acid molecule which is capable of regulating transcription of a human gene; and

(d) a further reporter nucleic acid molecule for producing a detectable amount of a further reporter molecule for indicating regulation of transcription of the further reporter nucleic acid molecule by the further regulatory nucleic acid molecule

wherein the further reporter and further regulatory nucleic acid molecules are arranged to permit the further

regulatory nucleic acid molecule to regulate transcription of the further reporter nucleic acid molecule.

In one embodiment, the at least one further regulatory nucleic acid molecule has a sequence shown in SEQ ID NO:4. In another embodiment, the at least one further regulatory nucleic acid molecule has a sequence shown in SEQ ID NO:5.

Although the regulatory nucleic acid molecule of the invention described herein is sufficient for reproducing the constitutive tissue specific and developmental expression of the CYP3A4 gene that is observed in a human, the inventors recognise that aspects of the xenobiotic inducibility of the gene could be better reproduced in an animal by incorporating at least one human transcription factor that is capable of interacting with the regulatory nucleic acid molecule for regulating transcription of the human CYP3A4 gene. Examples of such factors are nuclear receptors. These receptors may be those capable of regulating CYP3A4 gene transcription in a human when the receptor is bound to a ligand, such as a xenobiotic or steroid. One example of such a receptor is the human PXR (pregnane X receptor, otherwise known as SXR (steroid and xenobiotic receptor)). Another suitable receptor is the human CAR (constitutive androstane receptor-β). Non-human animals comprising a human PXR or CAR receptor are known. The inventors recognise that the genetic background of these animals could be incorporated into the non-human mammal of the present invention, for example, by conventional breeding techniques.

Thus in a sixth aspect, the non-human animal of the invention further comprises at least one human transcription factor for regulating transcription of a human CYP3A4 gene. Preferably the transcription factor is a nuclear receptor. Preferably, the nuclear receptor is a

heterodimer of the human PXR (pregnane X receptor, otherwise known as SXR (steroid and xenobiotic receptor)) and human RXR (9-cis retinoic acid receptor) or human CAR (constitutive androstane receptor-β) and human RXR.

It follows that the reporter nucleic acid molecule can be any molecule which is capable of detection when the reporter nucleic acid molecule is transcribed. For example, the reporter nucleic acid molecule could be the CYP3A4 cytochrome, or the mRNA transcript which is translated to produce the cytochrome. Those reporter molecules which are commercially available, including firefly luciferase, β- galactosidase, alkaline phosphatase, green fluorescent protein or chloramphenicol acetyl transferase can be used.

Thus in one embodiment, the reporter nucleic acid molecule is capable of producing a reporter molecule selected from the group of reporter molecules consisting of firefly luciferase, β-galactosidase, alkaline phosphatase, green fluorescent protein or chloramphenicol acetyl transferase.

While the non-human mammal of the invention, as exemplified below, is a mouse, the inventors believe that any other non-human mammal could be used in the invention, especially those for which standard transgenic techniques have been developed including for example, rat and rabbit. However, typically the non-human mammal is a mouse.

In another aspect, the invention provides a tissue of a non-human mammal of the invention.

In one embodiment, the tissue is an embryo capable of producing a non-human mammal of the invention.

In a further aspect, the invention provides a method of determining whether a compound is capable of effecting the transcription of a human CYP3A4 gene the method comprising the following steps:

(a) administering the compound to a non human mammal according to the invention and
(b) determining whether the reporter molecule is produced by the reporter nucleic acid molecule in the mammal.

In one embodiment, the production of the reporter molecule indicates that the binding compound is capable of effecting the transcription of the human CYP3A4 gene.

Any compound can be tested in the method however, preferred compounds are xenobiotic or steroid compounds.

The inventors recognise that a non human animal which comprises a 5' flanking region of CYP3A4 gene but which is deficient for the region from -7836 to -7207 would be useful as a negative control in a method for determining whether a compound is capable of regulating transcription of the human CYP3A4 gene.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1. CYP3A4/*lacZ* transgene constructs used to generate transgenic mice. The upstream regions of the human *CYP3A4* gene are depicted as open boxes with the position of the XREM at approximately -7.5kb of the CYP3A4 gene indicated by cross-hatching. The 5'-flanking region extended from 56bp downstream of the transcription initiation site to a HindIII site at -3,213 in the construct designated -3CYP3A4/*lacZ* and to a KpnI site at -12,926 kb in construct

-13CYP3A4/lacZ. The coding region of the E.coli *lacZ* gene together with eukaryotic translational initiation and termination signals, transcription termination and poly adenylation sites are indicated by a solid box.

Figure 2. Xenobiotic induction of hepatic transgene expression. Female mice from line 9/4 harbouring the -13CYP3A4/*lacZ* transgene were treated with various reagents. Histochemical staining of liver slices with X-gal revealed an increased zone of blue staining cells containing β-galactosidase after treatment with rifampicin, phenobarbital and pregnenolone 16α-carbonitrile compared with corn oil treated mice.

Figure 3. Comparison of the xenobiotic induction profile of the -13CYP3A4/*lacZ* transgene with the mouse Cyp3a11 gene. Transgenic mice from line 9/4 were treated with a range of xenobiotic reagents and naturally occurring steroids. **A.** Transgene expression was assessed by determining β-galactosidase activity in total liver lysates using the ONPG assay. The units of β-galactosidase activity are given as A_{420}/mg liver/minute. Dexamethasone and pregnenolone 16α-carbonitrile were the most potent xenobiotic activators of the -13CYP3A4/*lacZ* transgene, while rifampicin treatment resulted in relatively low levels. The steroids pregnenolone and 17α-progesterone were very weak inducers. **B.** Hepatic expression of the endogenous mouse Cyp3a11 gene was examined in the same mice by Northern analysis. A similar pattern of induction to the CYP3A4/*lacZ* transgene was observed with both xenobiotic and endogenous regulators. The data are presented as the mean +/- the standard deviation for 3 animals.

Figure 4. Dose response of -13CYP3A4/*lacZ* transgene expression after treatment with dexamethasone. **A.** Male mice from line 9/4 were treated with from 1 to 100mg/kg dexamethasone. Higher doses of dexamethasone resulted in increased β-galactosidase activity (determined in liver lysates as described in Fig. 3).**B.** Zonal expansion of transgene expression with increasing doses of dexamethasone. X-gal staining of frozen liver sections revealed greater numbers of hepatocytes containing transgene-derived β-galactosidase activity after treatment with 1, 10 and 100 mg/kg dexamethasone. At low doses there are limited numbers of transgene -expressing cells immediately adjacent to the central vein. With higher doses there are more cells committed to transgene expression extending across the liver lobule towards the portal tract.

Figure 5. (SEQ ID NO:1) Sequence of the CYP3A4 5'-flanking region included in the -13 CYP3A4/lacZ construct. This sequence corresponds to -12,926 to +56 base pairs relative to the transcription initiation site of the CYP3A4 gene.

Figure 6. (SEQ ID NO:2)Sequence of the 5'-flanking region of the CYP3A4 gene extending from -12,926 to -3,213 base pairs and representing the difference in sequence between the -13 CYP3A4/lacZ and the -3 CYP3A4/lacZ constructs.

Figure 7. (SEQ ID NO:3)The "Xenobiotic-Responsive Enhancer Module" (XREM) of the human CYP3A4 gene. This region encompasses -7836 to -7207 base pairs relative to the transcription initiation site of the CYP3A4 gene.

Figure 8. (SEQ ID NO:4)The 5'-flanking region of the human CYP3A7 gene (Genbank Accession No. AF329900). The extent of the sequences is -11,133 to +52 base relative to the transcription initiation site of the CYP3A7 gene.

Figure 9. (SEQ ID NO:5)Sequence of the 5'-flanking region of the human MDR1 gene (p-glycoprotein gene) encompassing -10,000 to +200 base pairs relative to the transcription initiation site of the MDR1 gene. Sequence derived from within Genbank sequence Accession Number AC002457.

An embodiment of the invention is now described in the following Example which will be understood to merely exemplify and not to limit the scope of the invention.

EXAMPLE

MATERIALS AND METHODS

Transgene constructs. Two transgene constructs were synthesized with the upstream 5' flank of the human cytochrome P450 *CYP3A4* gene linked to the E. coli *lacZ* reporter gene (Figure 1). The first construct, designated -3CYP3A4/*lacZ*, contained the region of the *CYP3A4* gene from the HindIII site at -3213bp relative to the transcription start site to nucleotide +56bp downstream of the transcription start site. The other construct, designated -13CYP3A4/*lacZ*, included the region of the CYP3A4 gene from the KpnI site at -12,926bp upstream to +56bp downstream of the transcription start site. It includes the DNA sequences of the XREM region located between -7836 and -7208 in addition to the proximal promoter of the *CYP3A4* gene. The DNA sequence of the *CYP3A4* gene between -10468bp and +906bp has been determined and deposited with the

GenBank/EMBL/DDJB database under accession number AF185589. Additional sequence information covering the region -10,469bp to -12,926bp was obtained from publically accessible Genbank files. The E.coli *lacZ* reporter gene comprises the coding region for the bacterial enzyme β-galactosidase flanked by DNA sequences for eukaryotic translational start and stop signals, SV40 transcriptional termination and polyadenylation signals and an intron. The *CYP3A4/lacZ* transgene constructs were released from vector sequences and purified on agarose gels prior to microinjection

Generation of transgenic mouse lines. Mice carrying the CYP3A4/*lacZ* transgenes were created by microinjection of the DNA constructs into the pro-nuclei of zygotes harvested from FVB/N strain mice. Microinjection and manipulation of embryos were carried by standard techniques. Stable transgenic mouse lines were established by breeding from transgenic founders identified by Southern analysis.

Administration of xenobiotics to mice. 8-10 week old male and female mice hemizygous for the -3CYP3A4/*lacZ* and -13CYP3A4/*lacZ* transgenes were used to test the ability of a range of xenobiotics and hormones to activate expression of transgene-derived β-galactosidase. Mice were administered the following reagents and vehicles by single daily intraperitoneal injection for 4 days: rifampicin/corn oil; dexamethasone phosphate/H_2O; pregnenolone 16α-carbonitrile/2% Tween 20 in H_2O; phenobarbital/H_2O; clotrimazole/2% Tween 20; phenytoin/2% Tween 20; 17α-OH progesterone/2% Tween 20; pregnenolone/2% Tween 20. All reagents were supplied by Sigma Chemical Co. (St Louis, MO) except for dexamethasone phosphate which was obtained from Faulding (Mulgrave, Australia) and pregnenolone 16α-

carbonitrile from Upjohn Co. (Kalamazoo, MI). The dose used for all reagents to test for induction of the transgene was 100mg/kg body weight. Dose response studies were carried out in the range of 1-100mg/kg with male hemizygous transgenic mice.

Analysis of transgene and mouse Cyp3a gene expression. β-galactosidase activity was visualised in slices and frozen sections of liver and other tissues by staining with X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside). Tissues were fixed in 0.25% glutaraldehyde, 0.1M phosphate buffer pH7.3, 5mM EGTA, 2 mM $MgCl_2$: washed in 0.1M phosphate buffer pH7.3, 0.01%sodium deoxycholate, 0.025% NP40, 2mM $MgCl_2$ and stained by incubation at 37°C in wash solution supplemented with 1mg/ml X-gal, 5mM potassium ferricyanide, and 5mM potassium ferrocyanide. The level of β-galactosidase activity was determined in whole liver homogenates [100mg fresh tissue/ml 0.25M Tris-HCl (pH 7.3)] using the O-nitrophenyl-β-D-galactopyranoside (ONPG) assay according to standard techniques. After appropriate dilution the homogenate was incubated with β-galactosidase assay reagent (0.1M sodium phosphate buffer (pH7.3)/1mM $MgCl_2$/50 mmol β-mercaptoethanol/0.88mg/ml ONPG) at 37°C, quenched by the addition of 1M Na_2CO_3 and the absorbance at 420nm determined. The units of β-galactosidase activity are given as A_{420}/mg liver/minute.

The levels of endogenous mouse Cyp3a mRNA expression were determined by Northern analysis using a riboprobe complementary to nucleotides 852-1061 of the mouse Cyp3a11 cDNA. Filters were stripped and reprobed with an 18S rRNA oligonucleotide to normalise loading.

RESULTS

4 transgenic lines were generated with the construct containing the -3.2kb region of the human *CYP3A4* gene linked to *lacZ*. Transgene-derived β-galactosidase activity was not detected in kidney, large and small intestine, spleen, lung and liver tissue from mice for all 4 -3CYP3A4/*lacZ* transgenic lines treated with vehicle or xenobiotics (Table 1). In contrast, transgene expression was readily detected in 3 of the 4 lines carrying the -13CYP3A4/*lacZ* construct. Line 9/4 had a very low constitutive level in the liver, with β-galactosidase detected only in isolated hepatocytes adjacent to major blood vessels. Administration of xenobiotics resulted in robust expression in a zone of cells surrounding the central vein (Figure 2). As the basal level of transgene expression in untreated mice in line 9/4 is extremely low, induction is obvious and is essentially an off/on process. Expression in other tissues in mice from line 9/4 was restricted to the gut, predominantly in the villi of the small intestine.

The relative degree of induction for a range of xenobiotics was analysed by determining the transgenic β-galactosidase activity in liver lysates of mice from line 9/4 (Figure 3A). Dexamethasone and pregnenolone 16α-carbonitrile were the most potent inducers, while rifampicin activated the transgene to relatively modest levels. Phenobarbital, clotrimazole and phenytoin were intermediate inducers. The induction profile of the transgene in line 9/4 was similar to that observed for the endogenous *Cyp3a11* gene in the same mice (Fig 3B), likely reflecting the activation profile of the mouse rather than the human PXR. Activation of the transgene was observed with naturally occurring

steroids such as pregnenolone and 17α-progesterone, however the induction was weak compared with xenobiotics.

There was a marked gender difference in hepatic transgene expression, with lower levels observed in females than in males for most reagents. Such a male-predominant pattern was not evident in the induction profile of the mouse *Cyp3a11* gene. Indeed higher levels of Cyp3a11 mRNA were observed in females than males after treatment with rifampicin and pregnenolone 16α-carbonitrile. The reason for this apparent reversal in gender-related transgene expression pattern is not known. However, as Cyp3a11 mRNA is only just detectable in males of the FVB/N strain of mice, it may be attributed to the relatively greater degree of induction of the mouse *Cyp3a11* gene in males compared to females (Figure 3B).

The other line which showed significant transgene expression - 15/10, had a higher constitutive level in both the liver and small intestine in untreated mice. Expression was not detected in other organs, confirming the tissue specificity observed in line 9/4. The same set of reagents were capable of increasing hepatic and intestinal transgene expression to the same levels as in mice from line 9/4. However, the overall degree of induction was not as great as observed in line 9/4 due to the higher basal level in line 15/10. The induction profile was similar with dexamethasone being the most potent activator and rifampicin the least (data not shown).

Dose response of xenobiotic induction. The activation of transgene expression in line 9/4 by dexamethasone was dose-dependent over the range 1 to 100 mg/kg (Figure 4A). The higher transgene-derived β-galactosidase activity in liver homogenates from mice treated with increasing doses of

dexamethasone was associated with an expanded zone of cells which were stained by X-gal. At low doses of dexamethasone a ring of hepatocytes only 1-2 cells thick around the central vein expressed the transgene (Figure 4B). With 100mg/kg dexamethasone the zone of X-gal positive hepatocytes increased to up to 10 cells, approximately midway between the central vein and the portal triad. A similar dose-dependent expansion of hepatocytes expressing the transgene was observed with other reagents and also in line 15/10 which also contained the -13CYP3A4/lacZ construct.

CLAIMS

1. A non-human mammal comprising:
(a) a regulatory nucleic acid molecule which is capable of regulating transcription of the human CYP3A4 gene and which comprises a nucleotide sequence that is identical to a sequence of the human CYP3A4 gene located between the initiation of transcription site of the gene and a position located at least 13,000 nucleotides upstream from the site; and
(b) a reporter nucleic acid molecule for producing a detectable amount of a reporter molecule for indicating regulation of transcription of the reporter nucleic acid molecule by the regulatory nucleic acid molecule;
wherein the reporter and regulatory nucleic acid molecules are arranged to permit the regulatory nucleic acid molecule to regulate transcription of the reporter nucleic acid molecule.

2. A non human mammal comprising:
(a) a regulatory nucleic acid molecule comprising a nucleotide sequence that is identical to the nucleotide sequence of the human CYP3A4 gene that extends about 13,000 nucleotides upstream from the initiation of transcription site of the gene; and
(b) a reporter nucleic acid molecule for producing a detectable amount of a reporter molecule for indicating regulation of transcription of the reporter nucleic acid molecule by the regulatory nucleic acid molecule;
wherein the reporter and regulatory nucleic acid molecules are arranged to permit the regulatory nucleic acid molecule to regulate transcription of the reporter nucleic acid molecule.

3. A mammal according to claim 2 wherein the regulatory nucleic acid molecule comprises the sequence shown in SEQ ID NO:1.

4. A non-human mammal comprising:
(a) a regulatory nucleic acid molecule comprising a nucleotide sequence that is identical to the sequence of the human CYP3A4 gene that extends about 8,000 nucleotides upstream from a position about 3,000 nucleotides upstream from the initiation of transcription site of the gene; and
(b) a reporter nucleic acid molecule for producing a detectable amount of a reporter molecule for indicating regulation of transcription of the reporter nucleic acid molecule by the regulatory nucleic acid molecule;
wherein the reporter and regulatory nucleic acid molecules are arranged to permit the regulatory nucleic acid molecule to regulate transcription of the reporter nucleic acid molecule.
5. A mammal according to claim 4 wherein the regulatory nucleic acid molecule comprises the sequence shown SEQ ID NO:2.
6. A non-human mammal comprising:
(a) a regulatory nucleic acid molecule which is capable of regulating transcription of the human CYP3A4 gene and which comprises a nucleotide sequence that is identical to the sequence of the human CYP3A4 gene that extends about 600 nucleotides upstream from a position about 7,200 nucleotides upstream of the initiation of transcription site of the gene; and
(b) a reporter nucleic acid molecule for producing a detectable amount of a reporter molecule for indicating regulation of transcription of the reporter nucleic acid molecule by the regulatory nucleic acid molecule;
wherein the reporter and regulatory nucleic acid molecules are arranged to permit the regulatory nucleic acid molecule to regulate transcription of the reporter nucleic acid molecule.

7. A mammal according to claim 6 wherein the regulatory nucleic acid molecule comprises the sequence shown in SEQ ID NO:3.

8. A mammal according to any one of the preceding claims wherein the regulatory nucleic acid molecule has the sequence of a fragment of the CYP3A4 gene consisting of from nucleotide positions −7836 to −7207 contiguous with -362 to +53.

9. A mammal according to any one of the preceding claims, further comprising:

(c) a further regulatory nucleic acid molecule which is capable of regulating transcription of a human gene; and

(d) a further reporter nucleic acid molecule for producing a detectable amount of a further reporter molecule for indicating regulation of transcription of the further reporter nucleic acid molecule by the further regulatory nucleic acid molecule;

wherein the further reporter and further regulatory nucleic acid molecules are arranged to permit the further regulatory nucleic acid molecule to regulate transcription of the further reporter nucleic acid molecule.

10. A mammal according to claim 9 wherein the at least one further regulatory nucleic acid molecule has a sequence shown in SEQ ID NO:4.

11. A mammal according to claim 9 wherein the at least one further regulatory nucleic acid molecule has a sequence shown in SEQ ID NO:5.

12. A mammal according to any one of the preceding claims, further comprising at least one human transcription factor for regulating transcription of a human CYP3A4 gene.

13. A mammal according to claim 12 wherein the transcription factor is a nuclear receptor.

14. A mammal according to claim 13 wherein the nuclear receptor is a heterodimer of the human pregnane X receptor and human 9-cis retinoic acid receptor or a heterodimer of human constitutive androstane receptor-β and human 9-cis retinoic acid receptor.

15. A mammal according to any one of the preceding claims wherein the reporter nucleic acid molecule is capable of producing a reporter molecule selected from the group of reporter molecules consisting of firefly luciferase, β-galactosidase, alkaline phosphatase, green fluorescent protein or chloramphenicol acetyl transferase.

16. A mammal according to any one of the preceding claims wherein the mammal is a mouse.

17. A tissue of a mammal according to any one of the preceding claims.

18. A tissue according to claim 17 wherein the tissue is an embryo capable of producing a mammal according to any one of the preceding claims.

19. A method of determining whether a compound is capable of effecting the transcription of a human CYP3A4 gene the method comprising the following steps:

    (a) administering the compound to a non human mammal according to any one of the preceding claims; and

    (b) determining whether the reporter molecule is produced by the reporter nucleic acid molecule in the mammal.

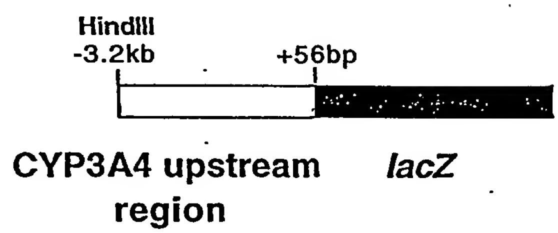
20. A method according to claim 19 wherein the production of the reporter molecule indicates that the binding compound is capable of effecting the transcription of the human CYP3A4 gene.

Table 1. Expression of CYP3A4/*lacZ* transgenic lines

Construct	Line No.	Copy No.	LIVER		Small Intestine
			Basal	Inducible	
-3CYP3A4/ lacZ	13	15	-	-	-
	24	>100	-	-	-
	31	80	-	-	-
	39	10	-	-	-
-13CYP3A4/ lacZ	13/5	70	-	-	
	9/4	5	+	++++	+
	9/7	50	-	+	-
	15/10	8	++	++++	++++

Fig 1. Human CYP3A4/*lacZ* transgene constructs

-3 CYP3A4/*lacZ*



-13 CYP3A4/*lacZ*

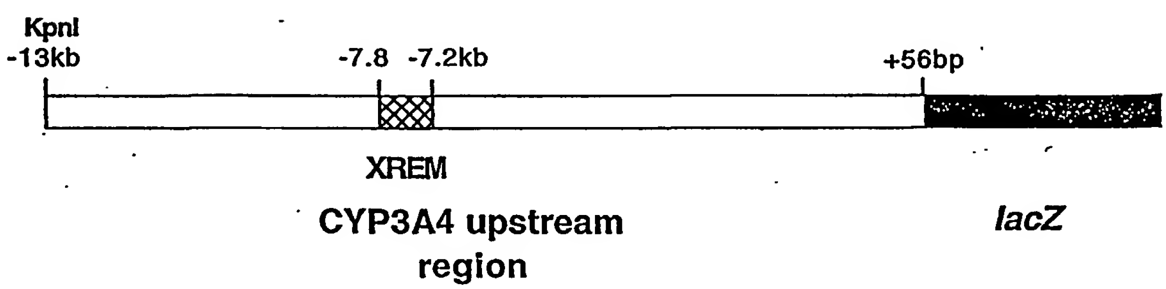


Figure 2. Xenobiotic induction of CYP3A4/lacZ transgene

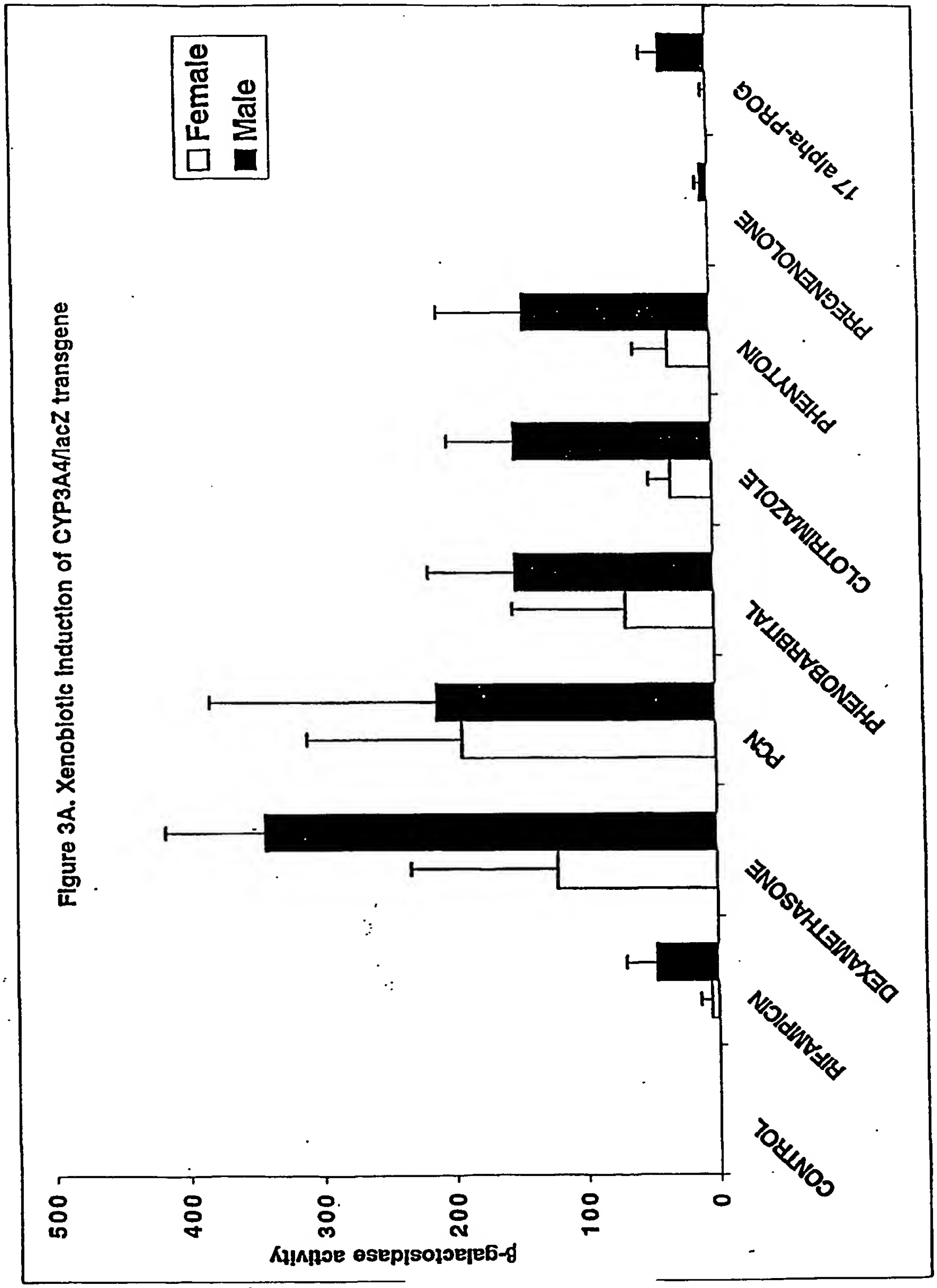


Figure 3A. Xenobiotic induction of CYP3A4/lacZ transgene

Fig 3B Xenobiotic induction of mouse CYP3A11 mRNA

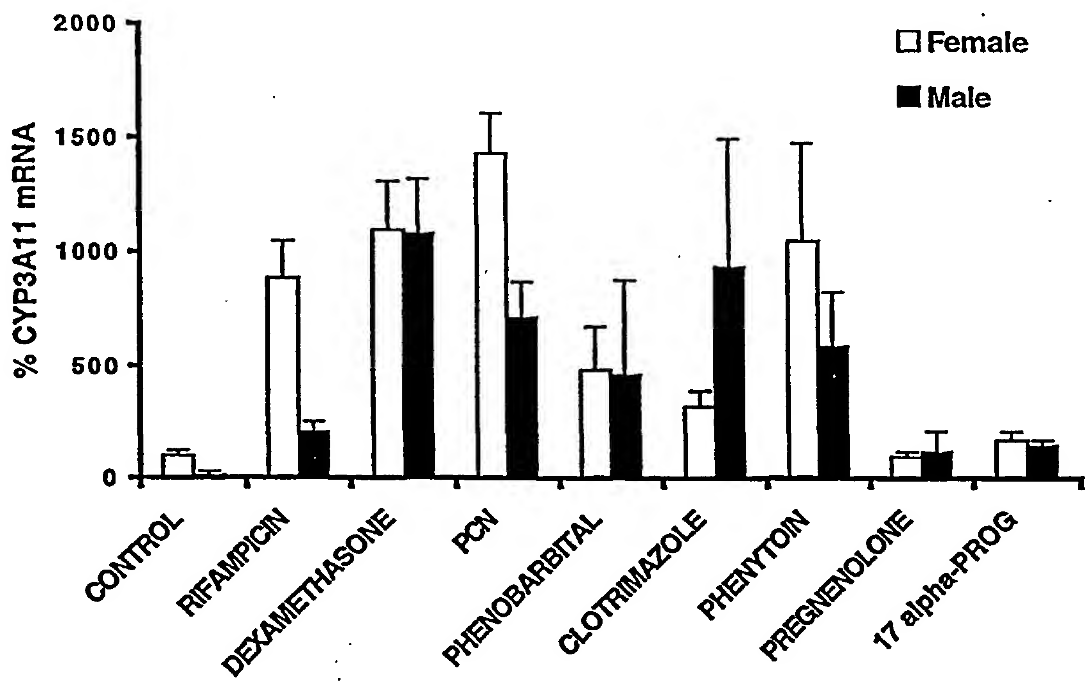


Figure 4A Dexamethasone dose-response (mg/kg body weight)

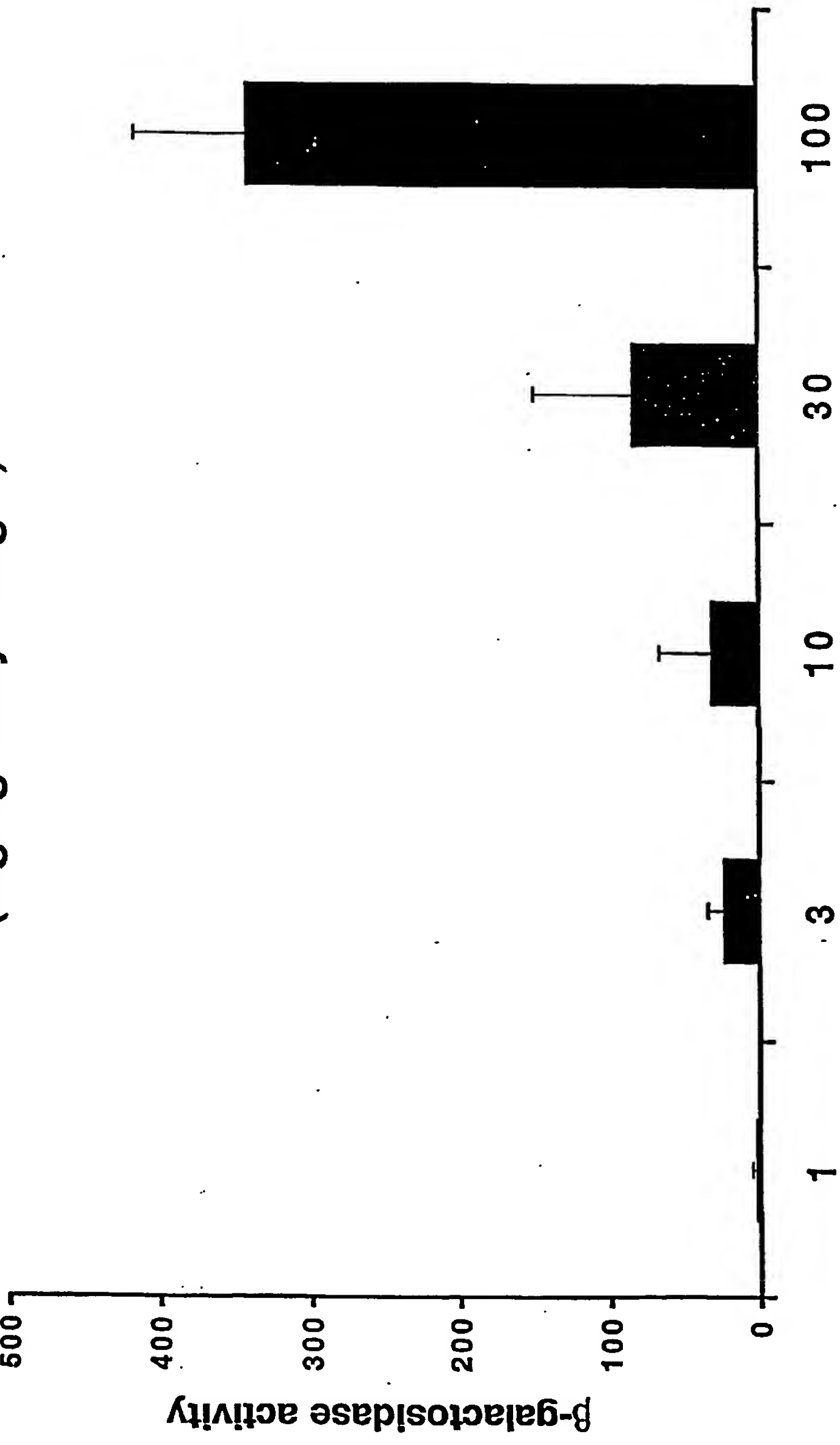
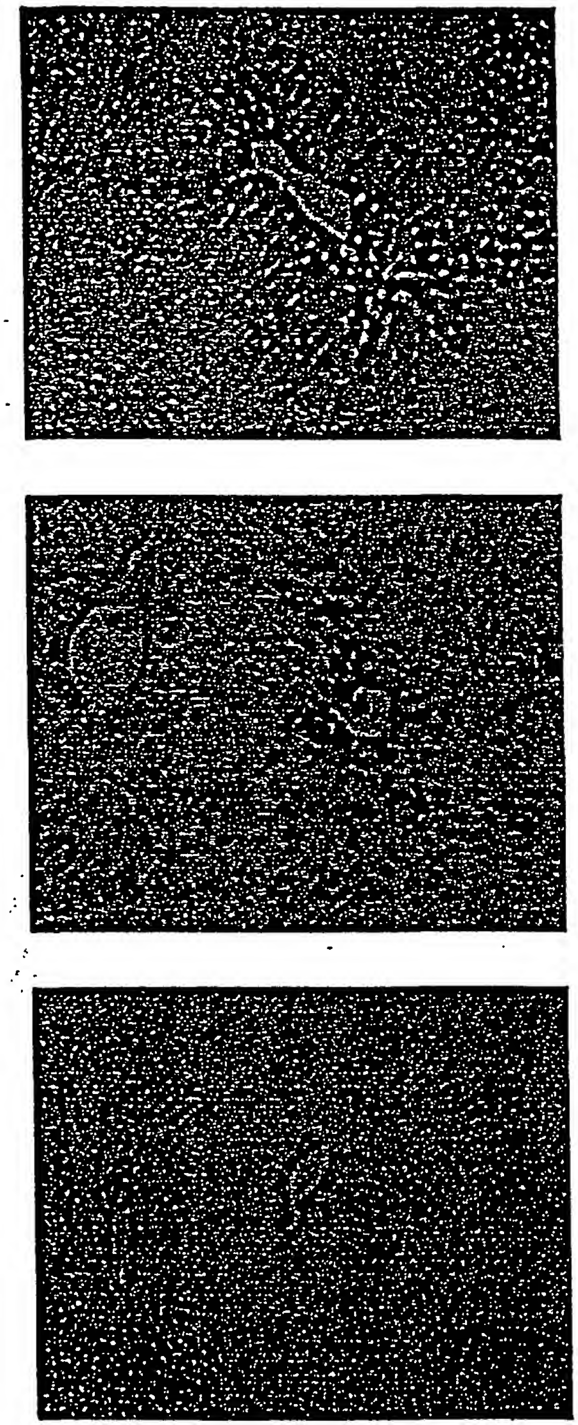


Figure 4B. Dose response of transgene expression



1 mg/kg

10 mg/kg

100 mg/kg

Figure 5.

CTGGTTCATCTCATTGGGACTGGTTGGACAAGAGGGTGCAGCCCACGGAGGGTGAGCCAAAGCAGGGTGGG
GCGTCGCCTCACCTGGGAAGCACAAGGGGTCGTGGAATTTTCTCCCCTACCCAAGGAAAGCCATAAGGGAC
TGAGCCTGAGGAACTGTGCACTCTGGCCCAGATACTGCACTTTTCCCATGGTCTTTGCAACCCGCAGACCA
GGAGATTCCCTCCGGTGCCTATGCCACCAGGGCCCTGGGTTTCAAGCACAAAACTGGGCAGCCATTTGGGC
AGACACCGAACTAGCTGCAGGAGTTTTTTTTTTTTTTTCCATACCCCATTGGCACCTGGAACGCCAGTGA
GACAGAACCGTTCACTCCCCTGGAAAGGGGGCTGAAACCAGGGATCCAAGTGGTCTGGCTCGGTGGCCCC
ACCCCCATGGAGCCCAGCAAACAAAGATTCACTTGGCTTGAAATTCTTGCTGCCAGCACAGCAGCAGTCTG
AGATTGACCTGGGACCCTCGAACTTGGTTGGGTGCTGTGGGGGGCATCTTCCATTGCTGAGGCTTGAGTA
GGTGGTTTTACCTTCGCGGTGTAAACAAAGCTGCTGGGAAGTTTGAACTGGGTGGAGCTCACCACAGCTCA
GTAAGGCCACTGTGGCCAGACTGCCTCTCTGGATTTCTCCTCTCTGGGAAGGATATCTCTGAAAAAAAGGC
AGCAGCCCCAGTCAGGGACTTATAGATGAAACCCCCATCTCCCTGGGACAGAGCCCCTCGGGGAAGAGGTG
GCTTCCACCATTGTGGAAGACTGTGTGGCAATTCCTCACGGATTTAGAACTAGAGATACCATTTGACCCAG
CAATCCCATTACTGGGTGTATACCCATAGGATTATAAATCATTCTACTATAAAGACACATGCACACTTATG
TTTATTGTAACACTATTTACAATAGCAATGACCTGGAACCAATCCAAAAGCCCATCAATGATAGACTGAAT
AAAGAAAATGTGGCACATATACACTGTGGAATACTATGCAGCCATAAAAAAGGATGAGTTCATGTCCTTTG
CAGAGACATGGATGAAGCTGGAAACCATCATTCTCAGCAAACTAGCACAATAACAGAAAACCAAACACTGC
ATGTTGTCACTCATAAGTGGGAGTTAAACAATGAGAACACATGGACACAGGGAGGGGAACGTCACACACTG
GGGCATGTCGGGGAGTGGGGGCCTACGGGAGGGATAGCATTAGCAGAAATACCTAATGTAGGTGACGGGTT
GATGGGTGCAGCAAACCACCATGGCACATATACACCTATGTAATAAAACTGCACGTTCTGCACATGTACCC
CAGAACTTAAAGTATAATTAATAATAATAATAATTTCTGGGCATGTAAGTAGCTGTCTTTCAGGTTCTACT
TTGATACATATTCTGAGAGAATTAAACCTGTCAAAGAAACCTTGACTTTCAATGGCAGGCACTGGAATTGA
CCCTAATAATGTGTTTTGGGGTAAGCCTACTCATATTCTCAACCTGTCTGCAGTAGTCGTTAGAATCTGAA
CTTCCTGAAGTTCATGTGCAAAGTTGAGTTAATTGTTTAATATTCAACAAGGATTATGCCAGTAAGATGGT
AGGAAAATATTAGATATGTGTCATCACTGCTGGTATTATTTAAACTGCAACATATTTTAGCTGGCTGCTGA
TCTCAGCCACCATGCCTGCATTTTATCTCTGTCTCGTGGTCTGCAACCTTGGAAGCTTTGAACTTAGCTCA
TAGAATCCTGGGCATCAAGAACATGTGGTTCTAATGGCTAGATAGGGAATGAGAGTAAAAGGATTTTGCCC
ACGGTCACGTGAGTAAACAACAGATTTGGAGGGGTCTGGACTACTGTGATGACTTCATTCTGACAATATGT
TCCAGTTGTCCTTTCATTTCCTCCTAATCACATGTCTGGTCTGATCTGGCTGTTTCCCACCTTCCAATTCC
TGCCTTCTCCAATGCTCCCTTCCGTAGGTCACTCTGTGGCTCAGAGACCCTGCTTAGCAAGCGCCCAACCT
TTCAATTATTTGTTCAGTAAAACTTGAACTCATGTCTCCCCTTCTTGATAAAAAGAAAATACGTTATGTAA
TGTCGGGTTACTCTATAACTCTTGTCCTGTCTCTCGGCAACTACTGAACTAACTGTTTTCATATTGAGCAA
ACGTTTATGGAAGGACTGCCAAGAGTCAGGTACTAGGCTTGGTAATATTCCCCGTTCTCTCTAGTCAAAGC
CAACACCAGCCAGACTTGCAGATCTAGGTCCCAAGCCCACTGCAGATCACAGGCCAGGGTCTGGTCTCCTC
TGAGCTCCTTTGGGAGGGAAAGACAGAATTATTAACACCCATTTTGTAGATTAGGCAACTGAGGCTGAGGA
AGTTTAAATAACTCAGACAGGGCCTGCACGTCAGTCATATTCCAAGGATCCCTACTCACTGTCTTCTCTCT
ACAGAACGAGATGTCTCTGGAGTCCATAGAAAGCCCAGGAGCCTGGCTGGGCACGGTGGCTCCTGCCTGTA
ATCCCAGCACTTTGGGAGGCCGAGGCAGGCAGATCACCTGAGCTCAGGAGTTCAAGACCAGCCTGGGCAAC
ATGGCAAAACCCCATCTCTACTAAAAATACAAAAAATTAGCTGGGCGTGGTGGTGCATGCCTCTAATCCCA

GCTACTTGGGAGGCTGAGGCACAAGAATTGCTTGAGCCCAGGAGGCAGCAGTTGCAGTGAGCTGAGATTGT
GCCAGTGCACTCCAGCCTGGGCAACAGAGCAAGATTCCATTTCAAAAACAAAAACAAACACAAACAAACAA
ACAAAAATAGAAAGCCCAGGGACCACCTGCGTCAGGTTCCCAGCCACACCTTTTTCTTGTCCTCCTCTGTC
TCTGGCATCTTCTCACAGGTTCCTAATTGTTTGTGGTTGCACAAATTCAAAATCCCAGAAAAATTACCACT
TCACACCCACTCAGATGGCTATTTTTTTTTTGAAGGAAGATAACAAGTGTTGACAAGAACATGGAGAAATT
GGAATTCTCACCCATTGCTGGTGAGAATGTAATACGGTGCTGCTGCTATGGAAAACAGCTTGGAGTTTCCT
CAAAAAGTTCAACAGAATTTCAATGTGACCCAGCAATTCCCCTCTAAGTTATAGATCTGAGAGGATTAAAA
ACAGTTACTAAAATACACGGACTCACATATTTCTAACAGTCCAATTCACAAGGGCCAAAAGGTGCTAATAG
CCCACATGTCCATCGATGGATGGATAAATAAATTGTGGTCTATCCATACAATGGAATATTATTCGGCCATA
AATGGAATGAAGTACTGACGCATGCTACAGAATGGATGAACCGCAAAAAAAATGGATGAACACATGCTACA
GAATGGATAGCCTCACTTTACTATGAAGTGAAGGCCAGAAACGAAGTCCATATATTGCATCATACAAAATA
TCCAGAAGAGGGAAGCCCACAGAGACAGAATGTGCAATGGTGGATGCCAGGGTCTGGGGAGAGGGGAGAGT
GGGGAGAAACTGCTCAACTGGTACAGGCTTTATTTTGGAATGATGGGAACATTTTGCAACTAGATAGAGGT
AGTGATTGCAGAACACAGAATGTACTGAATTCCACTGATTTTTTTCACCTTAAAATGGTTAATTTTCAGTC
CTGAGATTGGATAATCATAAAAAAAATGGTTAATTTTATGTTATGTGAATTTCATCCCTATACATATTTTAA
ACCTCAGAAATATACACTAGCAGGCATGGAACAGGTCACTGTGGTGCCTGCCAAGCCCGGTGATGTTATCT
GGGGTCCCCGGCCAGCCTTAAGCCTCTTGCTGACCGGTGGAGGGCAGAACCTTTGCCCTAAAAGTATAATA
TCCACATGCTGGCATGATTCCTGGCCAGATGGCTTCTTTATTAGCAGTAATTGAAACTGCCTCGATACAGA
CACTGTACCTTGCAACCAAAAAATGACTCAACAATGATAATAAGGGTTAAGCTGGGCCTTTCTCTCTTTGC
CAGTTAAATTATATTTATTATAGCTTGACATGAAAAACAAAGCAACTCCAACAGGTATCACAAGGGCAAAG
GACATGAACATTTTATCAAAGAAGAAATGCAGCTGTCAAAAATACAGAAATATTCAACCTTGTTCATAATA
AAGTGGCTGGGCTCAGTGGTTCATGCCTGTAATCCCAGTGCTTTGCAAGGCTGAGACAGGAGGATCATTTG
AAGCCAGAAGTTCAAGACCATCCTAGGCAAGTCAGTTCAATACCAGACTTCATGTCTACAAAACATCAAAA
AATTAGCCAGGCATGGTGATGCATGCCTGTTGTCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATTGCTTG
AGCCTGGGAGGCTGCGGTGGCGGTGAGCCATGATTGTGCCATTGTACTCCAGCCTGGGCAATGCAGCAAGA
CTGTCTAAATAACAAAAATAATAGTAAAGAAAAGGATTGGGATGCCATTTACTTGCGTATTCAATACACAG
AGTTAAAAGTAATTTCTACGTTTTCTATTTTTTTATTACTAAAAAAAGCTGGACCATTCTCACAGCCTGAA
ATGCTTCTCACTTTCCCTTCTTCTGTCCAAACACTTCTCTATGATAATGCAAACAGTCACTCCTTTAGGAA
GACTTCACCCCAGGTAGTTCCAGATCCCCTTATCTCTGCCTTCCCAGAACTCCTGGTGTCTCTCCAGTTCC
CTCCGTGTGGTGAAGTACCCTACCTAGGGTTTCAGTATGGCTCTGTCTGCAAAGGTCTTGTTCACACCTTC
CCTTATGGTTCTGTTGCCCTGTGTTGTGTCATAGCACAGGGCACAGTGGAGAACCCATTCACACTGATAGA
GAGGGCCCCATGGTCCTGGAGATAACCATGTAACCGATCAGAATAAGGCATTGAGGGCTGGGTGTCAGGCG
TGGGCTGCACTTGGGTGGGCAGGTCCCCTGGAAAGTCACTGGGTTTGGCAAGCTTCCTAGTAACATGTCTC
TCTGGGGTCCCCCTTGGAACTTCATGCAAAAATGCTGGTTGCTGGTTTATTCTAGAGAGATGGTTCATTCC
TTTCATTTGATTATCAAAGAAACTCATGTCCCAATTAAAGGTCATAAAGCCCAGTTTGTAAACTGAGATGA
TCTCAGCTGAATGAACTTGCTGACCCTCTGCTTTCCTCCAGCCTCTCGGTGCCCTTGAAATCATGTCGGTT
CAAGCAGCCTCATGAGGCATTACAAAGTTTAATTATTTCAGTGATTATTAAACCTTGTCCTGTGTTGACCC
CAGGTGAATCACAAGCTGAACTTCTGACAAGAACAAGCTATCATATTCTTTTCAATTACAGAAAAAAGTAA
GTTAATTGATAGGATTTTTTTTGTTTAAAAAAAATGTTACTAGTTTTGAAAAGGTAATATGTGCACATGGT
AAACACTAAGAAGGTATAAGAGCATAATGCTTTTATACTACTAAGAATAATGTTTTCTCTAAGTTTTTTTT

GGTAGATGCTTTCATCAGATTAAGAAAATTCCCTGCTATTAGTTGTTGAAGGTTTTTATATCATAAATGAA
AGTTGAATATTATTATCATATATTATTAATATATTGTTATTGAACTATCAAAGCCTTTTCCTAAAACCATT
GAGATGATCTTATAACCATTCTCCTTTAACCTGTTGACGAGATCATTGGTATTTATACTATTTCTCTGTTA
ACCATTCTTGAGTCTCAGGTTTAAATTCAACTTGGTCATGGTGTGTCATCTTTGATCATTGCTGTCTGTGG
CTTGCTACTGTTTTGTTTAGGATTTTTGCACTGATGCTCATCAATGAGACTGGCATGCCATCTTCCTTTGC
AGTCCTGATTTTTTTCTGATTTGGATCATGTGGTTATGGCCCTCATGGAATGAGTTGGGCATGATGCCTTT
TTTTCATGTCTCTGGATTGATGGGACACTTTGGATTCTCTCCAGATGGCCCTCAATGGTCCCTGCCTCCTC
ATTGTTAGGCCCCTGGGCAAGCCCTTCTCATTTCTGGTAGGCCCAGGAACCTGTGGGGGTTTTGTTTGTTT
GTTTGTTTCTTGAGTCGGAGTCTCACTCTGTCACCCAGGCTGGAGTTGGAGTGCAATGGCCCGATCTTGGC
TCACTGCAACCTCCACCTCCCAGATTCAAGCAATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGAATTACAG
GCACCCACCGACACACCCTGCTAATTTTTGTATTTTTAGTACAGATGGGGTTTCACAATATTGGCCAAGCT
GGTCTCGAACTCCTGATCTCATGATCTGCCCGGCTTGGCCTCCCAAAGTGTTGAGATTACAAGCATGAGCC
ACCACACCCAGTGAACCTGTGGTTTTTAGAAGCTCCCCATGCATGTGAATGCTGTGAGCATCCCAGGATGA
CAGCCACTGTGTGTTCAGCTGTTGGAACTGTGAGAAAGCACCAGTGGGACCTTCTCCAGCACCTGCCTGCT
GAGTTCATGGAAGAGGCTTGTTGGGGAGATGATGCCCTGGCTGACTCCTGAAGGATGGTTAGGAATGCACC
AGATGGAAGCTGGGTTGGACCCACTCTATGCTGAAGAACAGCTTGTGTGGACACAAGGAGACACGGATATG
TCATTTTTGTAGAGCCTGAGGAGTGTCCAATCACACCATTTGCTTAAAACATCATGCACACTTGGAAAAGT
GGACTGAGACCGAATGAAGAAGCTAACAGTGGCCAGATCAGAAAGGGTCTTGTGTTACTTCCTAGAGATAC
TTAGATTTTATCCTGTGGGTGATAGGAGCAGTTGGAGGGACTGAAGACAAGGAAAGAAACATGTTTCAAGA
TCTATGTTTTTCAAGACGCTTTTCTGGTGGCTGAGTAGGGAATTCCCTGGATAAGTCCTGCCCAGGGTCAG
GCAAAACAAGTTAGGGGGTTACTGAAATAAGGAGTATGAGAAATGGTGTAGGTTGTGCTGACGTTTTGTAA
CACATCTCATGATGATCTTCATTTCCTTCACTAATTTCCTGTTTCATTAATTCCCTTCCACGTGCTCTTCT
GAAATTTGCCTCACATTCTCTGATTTCTCTTTTACCTGTTGGTTTCATCACCTTTTACTTTTTGCTTTCCT
GGAAACACAAATGATTCTGATTGTGACATGTCAGAATTATTTGCAACATTTGCCTTTCTGCTGAAACCATG
AGTTCACTGAATACACAATTTAGTAAAGTGTAGGATGCACATGTCGTTTTCGTGGTCACAACCAGCTCTGT
AGCATTTTATAACTACACTGGCAGTGTGCTGGGAGGTGTAGAGAGAAATATTTATCACATGTGTGGCTGAC
ACAACCTGCCAAGTTATTTTAGGAGCCTCCTTGGAATCCCAGCAAGAATGCTACCGGCACAATTTGTAATC
ACAGCATCCTGCTCCATGCCTTGGCTTCATGGCATAGTCACTTCTGCAAGTCTCTTTCCAGCTGTCTGTTC
CCATGTCTATAAAGTATGAGTTAAATCATCCTAACACTACTCATCTTACAAAGTTTTCTTGCTGATGTTAA
GAGAGTTGGGAAAGAACTGTATAAACTGTGAAGTGCCATGGAGATGTTAGTGGTTACTTTATCAAGAAATA
GACACTCTAGAATGGAGTAGAAAGCCAACAGTTATGATTGAGTCCTCCTCCTCTTCTTCTTTTTATTAATT
TATAAAGAAAAGAGGTTTAATTGACTCACAGTTCCATATGGCTGGGGAGGCCTCGGGAAACTCTCAGTCAT
AGCAGGAGGCAAAGGGGAAGAAGGCACCTTCTTCACAAGGCGGCAGGAGAGAGAGAGCTCCTGTTCTTTTT
TGTCATAAAGTCTACAGAAGTGCTTATACTTCAGGACAAGGGCAGGCAGAGAGAAGGAAGGACATTGCTTC
ACCCCAGCCCTCACTGACGAGTTTGCTAGGGGACCTCACTTTGTCCCAGAGTAGGGCAGAACTCTGGCCAC
TACCCATTCAGAAGGCCTGGGCTGCACTGCTAGTTCCTCACTAACTCTGTGTGGCCTTGGGCAAGGTTGGG
CCTGTGTTAACAGATTATGACCCTGGGCTCTCAAGCTAGAGGATCTAAATTTGAATCCTGGCTCTGCTAAA
GCAATTAGTGATGTAAACTTTAATGGGTCAGTTAACCTTCCTGTGGCTTAGTTTGCTCATCTGTAAAATAG
GGATCATAACAGTATCAATACCACATGATTGTTGGACAGATTGAATCAGTTAATGCAGGGGAAGTACTTAG
CATGACACGTATTCACTATCATTTCCTGGAGTAAGAGCTGTGTGTGAGTGGGTGTGAGCATGTGTGAAACC

TTTTCTCTGCAATCTCAGTTAAGAAACCAATCCAGAATTTAAAGTTCAGGGCCTAAATGGGTGGTTATCTT
CTCCCAGTTCCATCCTATCCCACCTTTGCTCTTCCTCCCGCCCACAGGAGCTGTTGGTCCTTGATTGGGCT
GGAAGACCTGGTGGACCCTAAGTGATCTATAAGAGGAGAATAGAGAACAGGGAATGTCTTCAAAAATCTAG
AGGGACACAGAGGCTGAGAGGCAGGCAGTCCTGCAGGGTCTTCTGATTGGGACAAGGAGAACCTTGGTCTT
CACAGGCCAATTCTGGTCAGTTTCCCCCATGGACAGATGAGGAAACAGGCCCAGGAATATCCAAGGTCTCA
CACTTCCCATCTGTCAAGTCTTGTTGATTCTGTTGTATTCATGTCTCTCAAAGGGAGATAGAGTTTAGGGA
AGAAAGAAGGATCAACTGTGTCTGATACCACTGGGAGCTTAAGTAAAGGGTTCTTTTACTTCATAGCATTT
ATCCCAATTTGTAATTCAGTATTATTTGTGTGGCTGTTTGGTGTCTCTTTCTCCTATATGAGTGCTAGCTT
CATAAGGGCAAGGATTTTGATTCTTTAATATTTAGTGCTTGCCACATGCCCTGAACACAGCAGGCATACAG
GCTAACCAACATACAGTGGCATGAAAGTCATGAAAGTGAGACACCTACCTCCTCCAGTGCCAAGAGAGCAT
AACCATGCACCTGTCACTCTCCTCAACACCACCCCCAAGCATGAGGCCCAAAAGCATTAGCTAATCCCCTC
CTCCAGCCACTAAAACTTAAAGGCCAGGTGTGGTGGCTCCCATCTGAAATCCCAGAACTTCAGGAGACAGC
AGCAGGAGGATCACTTGAGGCCAGGAGTTTGAGATCAGCCTGGGCAACATAGCTAGGTCCCATCTGTACTA
AAAATTAGCTGGGCGTTGTTGCATGCCTGTAGTCCCAGCTACTAAGGAGGCTGAGGTGGGAGGATCACTTG
AGCCCAGGAGGTGGAAACAACAGTAAGCTATAATCACAGCACTGAACTCTAGCCTGGGCAACAGAGTGACA
CCCTGCCTCAAAACAATTTTAAAAATAAATAAGAGCAAAACTTAGATACCACGTGGTCACCCCAACATGCA
AAATCAAGTTTTCCCCTACTGAGAAGAATGGGGACTTGACAGCTGAGTTACAGAGAGATAATCTTCTTCTT
CTTTTTTTTTTTTTTGGTTTACATCCTCAAGATCATGACTTGTGAAATTTGAATCGAATACACATGTAATTC
CAGAGCAATGTTGCCTCCGCATACCATCAGCAATTCACTTGGCTACTGGAAGTCAGGATAAGCTTCCCAGA
AGAGAGGTACCACTTGGGCTACCAATATAAAAGGATGAAAATATCAGAGTGATGGTGTTCTTTACAACGTT
GAGTCCCTGGACAGCCTGTCCACTGATGCTGATATCTGAGCCTAATGCTTCTCTGAATGTTGAGATTGAAC
TTTGATCCAATGAAACTAGAACGAGAAAGAAGATAAGTCTTTCATTGTTGATAAGGACATTATGTTTCTCA
TACTTGTATGATTATTTTTCCTTAGCTGTACTATAATTATCTGCTTATTTGTCTCTGCTCTATGTGCTTAG
GGTACAAAGTTGACCAAGACCAACTTTGGTTGGAAGCATAGTACTAAGAGCACAGTACTGAGAGCACAGTA
TTGAGAGCACAGCTTTAAAAAACATGATGAAGGCTTTAATACAGGAAATGAGCAGGGGAGAGGCATGTGGT
GGTTGGATGTATCTTCCTTGACACAGTCAGTGCAGCTCTCAGTAGTCAAGTCCCTACATGTTAGAAGATGT
TACCTTCTGTGGAATTAAGTGGCAGAACTTGCCTTCAATTATTTTCCTTTGCAGAACAACACCAACTGCAT
TAGTTAGGACACAGTGCTGGCTGCATTTAAGTCCCAAGCGATGATTAGTCTCTCACTGTTGGTATAGATTC
AAACCAATCAGACCACCTCCTAAAGTTTGTAGGGCAGGTAAATCCTCATCTTAGAATAAAAATCATCTTAC
CAAGTATGTGTTTTAGAGGCAAGAAGAAAACATATTTGTTTCTGTAAGAGTTTTGTTTAAAAAAAATATAA
GAAAGGCTCTCGGTTTAGGTGAGGTAATGAAGTTGTTGATAGTTATCAGATGACACTGGAATCTTTACTTC
TCTGAACGTGTTCTGTGCATCTCTCAGTGTGGGAACATAGAGAGGGAGATCCTCCAGCAATGCCACTGATA
TGGTCAGAAACTGCATCTTTCTTTCTCCCTGCTGAGATGAGATGGAGTCCTTTGTTCTAGAAGACCCATGG
TGGTGCCGCTGGGAGTAACCCTTGAGACAGGAACACAAATCCCAACCAATTTGTGGTTGCAGCCTTGAGTC
TCACTATTTCCCATAGTGATGCGTAGCAGGGAATGGCAGGTGCACCAGAGCAGGAGAGGACCTAATATCTC
CCTTCCTGTTAGCTTTTTATAAAGTTTTATTGTGATCAGTAGCAGTTGGGAAGCTACTTGCAGTCACTGAG
CCTCAGTTTCTACATCTGTAAACTGGGGATAGTAGCATGGCCCCTACTTAATGTGCTCAGCAAAGCCACTG
AAAGGAGACAGAAATGTATCTAAATTACCCTGGACTTTTATCCTACCTCTCTTGGGGATTGTCACCACCTT
CCCATGTTTGTCCTTTTTGGTTTGATGCTTGCTGTCACTTCTTTCCTTAGGTGCCTCTCTGTACGGCTCTT
TTATCCCAGGGATTCCAGAGTTACAGCACATGCATACCACCATCCAAGCATGTTTATTTGTCTCCTGCTTC

ACTAGGCTGTCCCCAAGGAACATGTGGCTCCCGGCACACACCTGGCACAACACTGCACATGACATTCACCC
ACTTGGCCTTGAATCTGACAAGGAATCTGGCATGATGTTCACCCACTCAGGCCAGGTGCCGAGCAGCCCTG
GAGGCTTAGGGGCCAGAGGGATGGGAAAAGGTGTCTTTCTGGGGTGAGTATCAGTTTCTGCAGGAGGGCTG
AATGTGAGAAAGAATAAAGAGAGAAGGAAGCGAACAAGCACAGCTTAAACATCGCCTATTTCTATTGAGTT
TTAAGAACGCTGTGATTTTGTTTGTCATGCAATCCATTCATCAGGCCAGGCAGACACAGAACTTGGGTGTG
AGTGACGATAATGAGCTGATATAATTTTCACACCCTCATCACTGAGATCTCTCCCATCAGGAATGGGTCAG
GGAGCTCACAGGTGGCAGCAACTGCTATTACAGGCCTCATCTCTACCAGCTCCTGGGGCCTGCCCTCCTCC
CATTAGAAAATCCTCCACTTGTCAAAAAGGAAGCCATTTGCTTTGAACTCCAATTCCACCCCCAAGAGGCT
GGGACCATCTTACTGGAGTCCTTGATGCTGTGTGACCTGCAGTGACCACTGCCCCATCATTGCTGGCTGAG
GTGGTTGGGGTCCATCTGGCTATCTGGGCAGCTGTTCTCTTCTCTCCTTTCTCTCCTGTTTCCAGACATGC
AGTATTTCCAGAGAGAAGGGGCCACTCTTTGGCAAAGAACCTGTCTAACTTGCTATCTATGGCAGGACCTT
TGAAGGGTTCACAGGAAGCAGCACAAATTGATACTATTCCACCAAGCCATCAGCTCCATCTCATCCATGCC
CTGTCTCTCCTTTAGGGGTCCCCTTGCCAACAGAATCACAGAGGACCAGCCTGAAAGTGCAGAGACAGCAG
CTGAGGCACAGCCAAGAGCTCTGGCTGTATTAATGACCTAAGAAGTCACCAGAAAGTCAGAAGGGATGACA
TGCAGAGGCCCAGCAATCTCAGCTAAGTCAACTCCACCAGCCTTTCTAGTTGCCCACTGTGTGTACAGCAC
CCTGGTAGGGACCAGAGCCATGACAGGGAATAAGACTAGACTATGCCCTTGAGGAGCTCACCTCTGTTCAG
GGAAACAGGCGTGGAAACACAATGGTGGTAAAGAGGAAAGAGGACAATAGGATTGCATGAAGGGGATGGAA
GGTGCCCAGGGGAGGAAATGGTTACATCTGTGTGAGGAGTTTGGTGAGGAAAGACTCTAAGAGAAGGCTCT
GTCTGTCTGGGTTTGGAAGGATGTGTAGGAGTCTTCTAGGGGGCACAGGCACACTCCAGGCATAGGTAAAG
ATCTGTAGGTGTGGCTTGTTGGGATGAATTTCAAGTATTTTGGAATGAGGACAGCCATAGAGACAAGGGCA
AGAGAGAGGCGATTTAATAGATTTTATGCCAATGGCTCCACTTGAGTTTCTGATAAGAACCCAGAACCCTT
GGACTCCCCAGTAACATTGATTGAGTTGTTTATGATACCTCATAGAATATGAACTCAAAGGAGGTCAGTGA
GTGGTGTGTGTGTGATTCTTTGCCAACTTCCAAGGTGGAGAAGCCTCTTCCAACTGCAGGCAGAGCACAGG
TGGCCCTGCTACTGGCTGCAGCTCCAGCCCTGCCTCCTTCTCTAGCATATAAACAATCCAACAGCCTCACT
GAATCACTGCTGTGCAGGGCAGGAAAGCTCCATGCACATAGCCCAGCAAAGAGCAACACAG

Figure 6.

CTGGTTCATCTCATTGGGACTGGTTGGACAAGAGGGTGCAGCCCACGGAGGGTGAGCCAAAGCAGGGTGGG
GCGTCGCCTCACCTGGGAAGCACAAGGGGTCGTGGAATTTTCTCCCCTACCCAAGGAAAGCCATAAGGGAC
TGAGCCTGAGGAACTGTGCACTCTGGCCCAGATACTGCACTTTTCCCATGGTCTTTGCAACCCGCAGACCA
GGAGATTCCCTCCGGTGCCTATGCCACCAGGGCCCTGGGTTTCAAGCACAAAACTGGGCAGCCATTTGGGC
AGACACCGAACTAGCTGCAGGAGTTTTTTTTTTTTTTTTCCATACCCCATTGGCACCTGGAACGCCAGTGA
GACAGAACCGTTCACTCCCCTGGAAAGGGGGCTGAAACCAGGGATCCAAGTGGTCTGGCTCGGTGGGCCCC
ACCCCCATGGAGCCCAGCAAACAAAGATTCACTTGGCTTGAAATTCTTGCTGCCAGCACAGCAGCAGTCTG
AGATTGACCTGGGACCCTCGAACTTGGTTGGGTGCTGTGGGGGGGCATCTTCCATTGCTGAGGCTTGAGTA
GGTGGTTTTACCTTCGCGGTGTAAACAAAGCTGCTGGGAAGTTTGAACTGGGTGGAGCTCACCACAGCTCA
GTAAGGCCACTGTGGCCAGACTGCCTCTCTGGATTTCTCCTCTCTGGGAAGGATATCTCTGAAAAAAAGGC
AGCAGCCCCAGTCAGGGACTTATAGATGAAACCCCCATCTCCCTGGGACAGAGCCCCTCGGGGAAGAGGTG
GCTTCCACCATTGTGGAAGACTGTGTGGCAATTCCTCACGGATTTAGAACTAGAGATACCATTTGACCCAG
CAATCCCATTACTGGGTGTATACCCATAGGATTATAAATCATTCTACTATAAAGACACATGCACACTTATG
TTTATTGTAACACTATTTACAATAGCAATGACCTGGAACCAATCCAAAAGCCCATCAATGATAGACTGAAT
AAAGAAAATGTGGCACATATACACTGTGGAATACTATGCAGCCATAAAAAAGGATGAGTTCATGTCCTTTG
CAGAGACATGGATGAAGCTGGAAACCATCATTCTCAGCAAACTAGCACAATAACAGAAAACCAAACACTGC
ATGTTGTCACTCATAAGTGGGAGTTAAACAATGAGAACACATGGACACAGGGAGGGGAACGTCACACACTG
GGGCATGTCGGGGAGTGGGGGCCTACGGGAGGGATAGCATTAGCAGAAATACCTAATGTAGGTGACGGGTT
GATGGGTGCAGCAAACCACCATGGCACATATACACCTATGTAATAAACTGCACGTTCTGCACATGTACCC
CAGAACTTAAAGTATAATTAATAATAATAATAATTTCTGGGCATGTAAGTAGCTGTCTTTCAGGTTCTACT
TTGATACATATTCTGAGAGAATTAAACCTGTCAAAGAAACCTTGACTTTCAATGGCAGGCACTGGAATTGA
CCCTAATAATGTGTTTTGGGGTAAGCCTACTCATATTCTCAACCTGTCTGCAGTAGTCGTTAGAATCTGAA
CTTCCTGAAGTTCATGTGCAAAGTTGAGTTAATTGTTTAATATTCAACAAGGATTATGCCAGTAAGATGGT
AGGAAAATATTAGATATGTGTCATCACTGCTGGTATTATTTAAACTGCAACATATTTTAGCTGGCTGCTGA
TCTCAGCCACCATGCCTGCATTTTATCTCTGTCTCGTGGTCTGCAACCTTGGAAGCTTTGAACTTAGCTCA
TAGAATCCTGGGCATCAAGAACATGTGGTTCTAATGGCTAGATAGGGAATGAGAGTAAAAGGATTTTGCCC
ACGGTCACGTGAGTAAACAACAGATTTGGAGGGGTCTGGACTACTGTGATGACTTCATTCTGACAATATGT
TCCAGTTGTCCTTTCATTTCCTCCTAATCACATGTCTGGTCTGATCTGGCTGTTTCCCACCTTCCAATTCC
TGCCTTCTCCAATGCTCCCTTCCGTAGGTCACTCTGTGGCTCAGAGACCCTGCTTAGCAAGCGCCCAACCT
TTCAATTATTTGTTCAGTAAAACTTGAACTCATGTCTCCCCTTCTTGATAAAAAGAAAATACGTTATGTAA
TGTCGGGTTACTCTATAACTCTTGTCCTGTCTCTCGGCAACTACTGAACTAACTGTTTTCATATTGAGCAA
ACGTTTATGGAAGGACTGCCAAGAGTCAGGTACTAGGCTTGGTAATATTCCCCGTTCTCTCTAGTCAAAGC
CAACACCAGCCAGACTTGCAGATCTAGGTCCCAAGCCCACTGCAGATCACAGGCCAGGGTCTGGTCTCCTC
TGAGCTCCTTTGGGAGGGAAAGACAGAATTATTAACACCCATTTTGTAGATTAGGCAACTGAGGCTGAGGA
AGTTTAAATAACTCAGACAGGGCCTGCACGTCAGTCATATTCCAAGGATCCCTACTCACTGTCTTCTCTCT
ACAGAACGAGATGTCTCTGGAGTCCATAGAAAGCCCAGGAGCCTGGCTGGGCACGGTGGCTCCTGCCTGTA
ATCCCAGCACTTTGGGAGGCCGAGGCAGGCAGATCACCTGAGCTCAGGAGTTCAAGACCAGCCTGGGCAAC
ATGGCAAAACCCCATCTCTACTAAAAATACAAAAAATTAGCTGGGCGTGGTGGTGCATGCCTCTAATCCCA
GCTACTTGGGAGGCTGAGGCACAAGAATTGCTTGAGCCCAGGAGGCAGCAGTTGCAGTGAGCTGAGATTGT

GCCAGTGCACTCCAGCCTGGGCAACAGAGCAAGATTCCATTTCAAAAACAAAAACAAACACAAACAAACAAACAAAAATAGAAAGCCCAGGGACCACCTGCGTCAGGTTCCCAGCCACACCTTTTTCTTGTCCTCCTCTGTCTCTGGCATCTTCTCACAGGTTCCTAATTGTTTGTGGTTGCACAAATTCAAAATCCCAGAAAAATTACCACTTCACACCCACTCAGATGGCTATTTTTTTTTTGAAGGAAGATAACAAGTGTTGACAAGAACATGGAGAAATTGGAATTCTCACCCATTGCTGGTGAGAATGTAATACGGTGCTGCTGCTATGGAAAACAGCTTGGAGTTTCCTCAAAAAGTTCAACAGAATTTCAATGTGACCCAGCAATTCCCCTCTAAGTTATAGATCTGAGAGGATTAAAAACAGTTACTAAAATACACGGACTCACATATTTCTAACAGTCCAATTCACAAGGGCCAAAAGGTGCTAATAGCCCACATGTCCATCGATGGATGGATAAATAAATTGTGGTCTATCCATACAATGGAATATTATTCGGCCATAAATGGAATGAAGTACTGACGCATGCTACAGAATGGATGAACCGCAAAAAAAATGGATGAACACATGCTACAGAATGGATAGCCTCACTTTACTATGAAGTGAAGGCCAGAAACGAAGTCCATATATTGCATCATACAAAATATCCAGAAGAGGGAAGCCCACAGAGACAGAATGTGCAATGGTGGATGCCAGGGTCTGGGGAGAGGGGAGAGTGGGGAGAAACTGCTCAACTGGTACAGGCTTTATTTTGGAATGATGGGAACATTTTGCAACTAGATAGAGGTAGTGATTGCAGAACACAGAATGTACTGAATTCCACTGATTTTTTTCACCTTAAAATGGTTAATTTTCAGTCCTGAGATTGGATAATCATAAAAAAATGGTTAATTTTATGTTATGTGAATTTCATCCCTATACATATTTTAAACCTCAGAAATATACACTAGCAGGCATGGAACAGGTCACTGTGGTGCCTGCCAAGCCCGGTGATGTTATCTGGGGTCCCCGGCCAGCCTTAAGCCTCTTGCTGACCGGTGGAGGGCAGAACCTTTGCCCTAAAAGTATAATATCCACATGCTGGCATGATTCCTGGCCAGATGGCTTCTTTATTAGCAGTAATTGAAACTGCCTCGATACAGACACTGTACCTTGCAACCAAAAAATGACTCAACAATGATAATAAGGGTTAAGCTGGGCCTTTCTCTCTTTGCCAGTTAAATTATATTTATTATAGCTTGACATGAAAAACAAAGCAACTCCAACAGGTATCACAAGGGCAAAGGACATGAACATTTTATCAAAGAAGAAATGCAGCTGTCAAAAATACAGAAATATTCAACCTTGTTCATAATAAAGTGGCTGGGCTCAGTGGTTCATGCCTGTAATCCCAGTGCTTTGCAAGGCTGAGACAGGAGGATCATTTGAAGCCAGAAGTTCAAGACCATCCTAGGCAAGTCAGTTCAATACCAGACTTCATGTCTACAAAACATCAAAAAATTAGCCAGGCATGGTGATGCATGCCTGTTGTCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATTGCTTGAGCCTGGGAGGCTGCGGTGGCGGTGAGCCATGATTGTGCCATTGTACTCCAGCCTGGGCAATGCAGCAAGACTGTCTAAATAACAAAAATAATAGTAAAGAAAAGGATTGGGATGCCATTTACTTGCGTATTCAATACACAGAGTTAAAAGTAATTTCTACGTTTTCTATTTTTTTATTACTAAAAAAAGCTGGACCATTCTCACAGCCTGAAATGCTTCTCACTTTCCCTTCTTCTGTCCAAACACTTCTCTATGATAATGCAAACAGTCACTCCTTTAGGAAGACTTCACCCCAGGTAGTTCCAGATCCCCTTATCTCTGCCTTCCCAGAACTCCTGGTGTCTCTCCAGTTCCCTCCGTGTGGTGAAGTACCCTACCTAGGGTTTCAGTATGGCTCTGTCTGCAAAGGTCTTGTTCACACCTTCCCTTATGGTTCTGTTGCCCTGTGTTGTGTCATAGCACAGGGCACAGTGGAGAACCCATTCACACTGATAGAGAGGGCCCCATGGTCCTGGAGATAACCATGTAACCGATCAGAATAAGGCATTGAGGGCTGGGTGTCAGGCGTGGGCTGCACTTGGGTGGGCAGGTCCCCTGGAAAGTCACTGGGTTTGGCAAGCTTCCTAGTAACATGTCTCTCTGGGGTCCCCCTTGGAACTTCATGCAAAAATGCTGGTTGCTGGTTTATTCTAGAGAGATGGTTCATTCCTTTCATTTGATTATCAAAGAAACTCATGTCCCAATTAAAGGTCATAAAGCCCAGTTTGTAAACTGAGATGATCTCAGCTGAATGAACTTGCTGACCCTCTGCTTTCCTCCAGCCTCTCGGTGCCCTTGAAATCATGTCGGTTCAAGCAGCCTCATGAGGCATTACAAAGTTTAATTATTTCAGTGATTATTAAACCTTGTCCTGTGTTGACCCCAGGTGAATCACAAGCTGAACTTCTGACAAGAACAAGCTATCATATTCTTTTCAATTACAGAAAAAAGTAAGTTAATTGATAGGATTTTTTTTGTTTAAAAAAAATGTTACTAGTTTTGAAAAGGTAATATGTGCACATGGTAAACACTAAGAAGGTATAAGAGCATAATGCTTTTATACTACTAAGAATAATGTTTTCTCTAAGTTTTTTTTGGTAGATGCTTTCATCAGATTAAGAAAATTCCCTGCTATTAGTTGTTGAAGGTTTTTATATCATAAATGAA

AGTTGAATATTATTATCATATATTATTAATATATTGTTATTGAACTATCAAAGCCTTTTCCTAAAACCATT
GAGATGATCTTATAACCATTCTCCTTTAACCTGTTGACGAGATCATTGGTATTTATACTATTTCTCTGTTA
ACCATTCTTGAGTCTCAGGTTTAAATTCAACTTGGTCATGGTGTGTCATCTTTGATCATTGCTGTCTGTGG
CTTGCTACTGTTTTGTTTAGGATTTTTGCACTGATGCTCATCAATGAGACTGGCATGCCATCTTCCTTTGC
AGTCCTGATTTTTTTCTGATTTGGATCATGTGGTTATGGCCCTCATGGAATGAGTTGGGCATGATGCCTTT
TTTTCATGTCTCTGGATTGATGGGACACTTTGGATTCTCTCCAGATGGCCCTCAATGGTCCCTGCCTCCTC
ATTGTTAGGCCCCTGGGCAAGCCCTTCTCATTTCTGGTAGGCCCAGGAACCTGTGGGGGTTTTGTTTGTTT
GTTTGTTTCTTGAGTCGGAGTCTCACTCTGTCACCCAGGCTGGAGTTGGAGTGCAATGGCCCGATCTTGGC
TCACTGCAACCTCCACCTCCCAGATTCAAGCAATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGAATTACAG
GCACCCACCGACACACCCTGCTAATTTTTGTATTTTTAGTACAGATGGGGTTTCACAATATTGGCCAAGCT
GGTCTCGAACTCCTGATCTCATGATCTGCCCGGCTTGGCCTCCCAAAGTGTTGAGATTACAAGCATGAGCC
ACCACACCCAGTGAACCTGTGGTTTTTAGAAGCTCCCCATGCATGTGAATGCTGTGAGCATCCCAGGATGA
CAGCCACTGTGTGTTCAGCTGTTGGAACTGTGAGAAAGCACCAGTGGGACCTTCTCCAGCACCTGCCTGCT
GAGTTCATGGAAGAGGCTTGTTGGGGAGATGATGCCCTGGCTGACTCCTGAAGGATGGTTAGGAATGCACC
AGATGGAAGCTGGGTTGGACCCACTCTATGCTGAAGAACAGCTTGTGTGGACACAAGGAGACACGGATATG
TCATTTTTGTAGAGCCTGAGGAGTGTCCAATCACACCATTTGCTTAAAACATCATGCACACTTGGAAAAGT
GGACTGAGACCGAATGAAGAAGCTAACAGTGGCCAGATCAGAAAGGGTCTTGTGTTACTTCCTAGAGATAC
TTAGATTTTATCCTGTGGGTGATAGGAGCAGTTGGAGGGACTGAAGACAAGGAAAGAAACATGTTTCAAGA
TCTATGTTTTTCAAGACGCTTTTCTGGTGGCTGAGTAGGGAATTCCCTGGATAAGTCCTGCCCAGGGTCAG
GCAAAACAAGTTAGGGGGTTACTGAAATAAGGAGTATGAGAAATGGTGTAGGTTGTGCTGACGTTTTGTAA
CACATCTCATGATGATCTTCATTTCCTTCACTAATTTCCTGTTTCATTAATTCCCTTCCACGTGCTCTTCT
GAAATTTGCCTCACATTCTCTGATTTCTCTTTTACCTGTTGGTTTCATCACCTTTTACTTTTTGCTTTCCT
GGAAACACAAATGATTCTGATTGTGACATGTCAGAATTATTTGCAACATTTGCCTTTCTGCTGAAACCATG
AGTTCACTGAATACACAATTTAGTAAAGTGTAGGATGCACATGTCGTTTTCGTGGTCACAACCAGCTCTGT
AGCATTTTATAACTACACTGGCAGTGTGCTGGGAGGTGTAGAGAGAAATATTTATCACATGTGTGGCTGAC
ACAACCTGCCAAGTTATTTTAGGAGCCTCCTTGGAATCCCAGCAAGAATGCTACCGGCACAATTTGTAATC
ACAGCATCCTGCTCCATGCCTTGGCTTCATGGCATAGTCACTTCTGCAAGTCTCTTTCCAGCTGTCTGTTC
CCATGTCTATAAAGTATGAGTTAAATCATCCTAACACTACTCATCTTACAAAGTTTTCTTGCTGATGTTAA
GAGAGTTGGGAAAGAACTGTATAAACTGTGAAGTGCCATGGAGATGTTAGTGGTTACTTTATCAAGAAATA
GACACTCTAGAATGGAGTAGAAAGCCAACAGTTATGATTGAGTCCTCCTCCTCTTCTTCTTTTTATTAATT
TATAAAGAAAAGAGGTTTAATTGACTCACAGTTCCATATGGCTGGGGAGGCCTCGGGAAACTCTCAGTCAT
AGCAGGAGGCAAAGGGGAAGAAGGCACCTTCTTCACAAGGCGGCAGGAGAGAGAGAGCTCCTGTTCTTTTT
TGTCATAAAGTCTACAGAAGTGCTTATACTTCAGGACAAGGGCAGGCAGAGAGAAGGAAGGACATTGCTTC
ACCCCAGCCCTCACTGACGAGTTTGCTAGGGGACCTCACTTTGTCCCAGAGTAGGGCAGAACTCTGGCCAC
TACCCATTCAGAAGGCCTGGGCTGCACTGCTAGTTCCTCACTAACTCTGTGTGGCCTTGGGCAAGGTTGGG
CCTGTGTTAACAGATTATGACCCTGGGCTCTCAAGCTAGAGGATCTAAATTTGAATCCTGGCTCTGCTAAA
GCAATTAGTGATGTAAACTTTAATGGGTCAGTTAACCTTCCTGTGGCTTAGTTTGCTCATCTGTAAAATAG
GGATCATAACAGTATCAATACCACATGATTGTTGGACAGATTGAATCAGTTAATGCAGGGGAAGTACTTAG
CATGACACGTATTCACTATCATTTCCTGGAGTAAGAGCTGTGTGTGAGTGGGTGTGAGCATGTGTGAAACC
TTTTCTCTGCAATCTCAGTTAAGAAACCAATCCAGAATTTAAAGTTCAGGGCCTAAATGGGTGGTTATCTT

CTCCCAGTTCCATCCTATCCCACCTTTGCTCTTCCTCCCGCCCACAGGAGCTGTTGGTCCTTGATTGGGCT
GGAAGACCTGGTGGACCCTAAGTGATCTATAAGAGGAGAATAGAGAACAGGGAATGTCTTCAAAAATCTAG
AGGGACACAGAGGCTGAGAGGCAGGCAGTCCTGCAGGGTCTTCTGATTGGGACAAGGAGAACCTTGGTCTT
CACAGGCCAATTCTGGTCAGTTTCCCCCATGGACAGATGAGGAAACAGGCCCAGGAATATCCAAGGTCTCA
CACTTCCCATCTGTCAAGTCTTGTTGATTCTGTTGTATTCATGTCTCTCAAAGGGAGATAGAGTTTAGGGA
AGAAAGAAGGATCAACTGTGTCTGATACCACTGGGAGCTTAAGTAAAGGGTTCTTTTACTTCATAGCATTT
ATCCCAATTTGTAATTCAGTATTATTTGTGTGGCTGTTTGGTGTCTCTTTCTCCTATATGAGTGCTAGCTT
CATAAGGGCAAGGATTTTGATTCTTTAATATTTAGTGCTTGCCACATGCCCTGAACACAGCAGGCATACAG
GCTAACCAACATACAGTGGCATGAAAGTCATGAAAGTGAGACACCTACCTCCTCCAGTGCCAAGAGAGCAT
AACCATGCACCTGTCACTCTCCTCAACACCACCCCCAAGCATGAGGCCCAAAAGCATTAGCTAATCCCCTC
CTCCAGCCACTAAAACTTAAAGGCCAGGTGTGGTGGCTCCCATCTGAAATCCCAGAACTTCAGGAGACAGC
AGCAGGAGGATCACTTGAGGCCAGGAGTTTGAGATCAGCCTGGGCAACATAGCTAGGTCCCATCTGTACTA
AAAATTAGCTGGGCGTTGTTGCATGCCTGTAGTCCCAGCTACTAAGGAGGCTGAGGTGGAGGATCACTTG
AGCCCAGGAGGTGGAAACAACAGTAAGCTATAATCACAGCACTGAACTCTAGCCTGGGCAACAGAGTGACA
CCCTGCCTCAAAACAATTTTAAAAATAAATAAGAGCAAAACTTAGATACCACGTGGTCACCCCAACATGCA
AAATCAAGTTTTCCCCTACTGAGAAGAATGGGGACTTGACAGCTGAGTTACAGAGAGATAATCTTCTTCTT
CTTTTTTTTTTTTTGGTTTACATCCTCAAGATCATGACTTGTGAAATTTGAATCGAATACACATGTAATTC
CAGAGCAATGTTGCCTCCGCATACCATCAGCAATTCACTTGGCTACTGGAAGTCAGGAT

Figure 7.

TCTAGAGAGA TGGTTCATTC CTTTCATTTG ATTATCAAAG AAACTCATGT CCCAATTAAA
GGTCATAAAG CCCAGTTTGT AAACTGAGAT GATCTCAGCT GAATGAACTT GCTGACCCTC
TGCTTTCCTC CAGCCTCTCG GTGCCCTTGA AATCATGTCG GTTCAAGCAG CCTCATGAGG
CATTACAAAG TTTAATTATT TCAGTGATTA TTAAACCTTG TCCTGTGTTG ACCCCAGGTG
AATCACAAGC TGAACTTCTG ACAAGAACAA GCTATCATAT TCTTTTCAAT TACAGAAAAA
AGTAAGTTAA TTGATAGGAT TTTTTTTGTT TAAAAAAAAT GTTACTAGTT TTTGAAAAGG
TAATATGTTG CACATGGTAA ACACTAAGAA GGTATAAGAG CATAATGCTT TTATACTACT
AAGAATAATG TTTTCTCTAA GTTTTTTTTG GTAGATGCTT TCATCAGATT AAGAAAATTC
CCTGCTATTA GTTGTTGAAG GTTTTTATAT CATAAATGAA AGTTGAATAT TATTATCATA
TATTATTAAT ATATTGTTAT TGAACTATCA AAGCCTTTTC CTAAAACCAT TGAGATGATC
TTATAACCAT TCTCCTTTAA CCTGTTGACG AG

Figure 8.

GGATCCAGTTTCAGCTTTCTACATATGGCTAGCCAGTTTTCCCAGCACCATTTATTAAATAGGGAATCCTT
TCCCCATTGCTTGTTTTTGTCAGGTTTGTCAAAGATCAGATGGTTGTAGATGTGTGGTGTTTGTTCTGAGG
CCTCTGTTCTGTTCCATTGGTCCATATCCCTGTTTTGGTACTAGTACCATGCTCTTTTGGTTACTGTAGCC
TTGTAGTATAGTTTGAAGTCAGGTAGCGTGATTCCTCCAGCTTTGCTCTTTTTGCTTAGGATTGTCTTGGG
AATGTGGGCTCTTTTTTGGTTCCATATGAAATTTAAAGTAGTTTTTTTTCCAATTCTATGAAGAAAGTCAT
TGGTAACTTGATGGGGATGGCATTGAATCTATAAATTACCTTGGGAAGTATGGCCATTTTCACGATATTGA
TTCTTCCTATCCATGAGCATGGAACATTCTTCCATTTGTTTGTGTCCTCTTTGATTTTGTTGAGCAGTGGT
TTGTAGTTCTCCTTGAAGAAGTCCTTCACCTCCCTTTAATTTGGATTACTAGATATTTTATTCTCTTAGTA
ACAATTGCAAATGGGAGTTCACTCATGATTTGGCTCTCTTTCTGTTATTGGTGTATAGGAATGCTTGTGAT
TTTTGCGCATTAATTTTGTATCCTGAGACTTTGCTGAAGTTGCTTATCAGCTTAAAAGGATTTTGGGCTGA
GACGATGGGGTTTTCTAAATATACAATCATGGCATCTGCAAACAGGAACAATTTGACTTCCTCTTTTCCTA
ATTGAATACCCTTTATTTCTTTTTCTTGCCTGATTGCCCTGGCCAGAACTTCCAATACTATGTTGAATAAG
AGTCATGAGTGAGGGCATCGTTGTCTTGTGCTGGTTTCAAAGTTTTTGCCCATTCAGTATGATTTTGGCTG
TGGTTTTGCCATAAATAGCTCTTATTATTTTGAGATACGTTCCACCAATACCTACTTTATTGAGAGTTTTT
AGCAGGAAGGGCTGTTGAATTTTGTCGAAGGCCTTTTCTACATCTATTGAGACAATTATGTGGTTTTTTAA
TCGTTGATTCTGTTTATGTGATGGATTACATTTATTAATTTGCATATGTTGAACCAGCCTTGCATCCCAGG
GATGAAGCCCACTTGATTGTAGTGGATAAGCTTTTTGATGTGCTGCTGGATTCAGTTTGCCAGTATTTTAT
TGAGGATTTTGGCATCAATGTTCATCAGGGATATTGGTCTAAAATTCTCTTTTTTTGTTGTGTCTCTGCCA
GGCTTTGGTATCAGGATGATGCAGGCCTCAGAAACTGAGTTAGGGAGGATTCCCTCATTTTCTATTGATTG
GAATAGTTTCAGAAAGAATGGTACCAGCTACTCTTTGTACCTCTGGTAGAATTCAGCTGTGAATCCATCTG
GTCCTGGACTTTTTGGTTGGTAGGCTATTAATTATTGCCTCAATTTTAGGGCCTGTTATTGGTCTATTCAG
ACATTCAACTTCTTCCCGGTTTGGTCTTGGGAGGGTTTATGTGTCCAGGAATTTATCCATTTCTTCTAGAT
TTTCTAGTTTATTTGTGTAGAGGTGTTTATAGTATTGTCTGATGGTAGTTTGTATTTCTGTGAGATCGGTG
GTGATATCCCCTTTATCATTTTTTATTGCATCTATTTAATTCTTCTCTCTTTTCTTCTTTATTATTCTGGC
TGGCGGTCTGTCAATTTTTTTGATCTTTTCAAAAAACCAGCTCCTGGGTTTCACTGATTATTTGAAGGGTT
TTTTGTGTCTCTATTTCTTTCAGTTCTCCTGTGATCTTAGTTATTTCTTGCCTTCTGCTAGCTTTTGAATG
TGTTTGCTCTTCCTTCTCTAGTTCTTTGAATTGTGATGTTACAGTGTTGATTTTAGATCTTTCCTGCTTTC
TCTTGTGGTCATTTAGTGCTATAAATTTCCCTCTACACATTGGTTTACATGTGTCTCAGAGATTCTGGTAT
GTTGTGTCTTTGTTCTCATTCATTTCAAGAACATCTTTACTTCTGCCTTCATTTTGTTATTTGCCCAGTAG
TCATTCAGGAGCAGGTTGTTCAGTCTTCATGTAGTTGTGTGGTTTTGAGTGAGTTTCTTAATCCTGAGTTC
TAATTTGATTGCACTGTTGTCTGAGAGACAGTTTGTTGTGATTTCCATTCTTTTACATTTACTGAGCATGC
TTTATGTCCCATTATGTGGTCAATTTTAGAATAAGTGTGATGTGATGCTGAGAAGAATGTATATTCTGTTG
ATTTGGGGTGTGGAGTTCTGTAGATGTCTATTCAGTCCACTGGGTGCAGAGCTGAGTGGACATGAACATTT
TATCAAAGAAGAAACACAGCTATCAAAAATCCAGAAATATTGAACCTTGTTAATAATAAAGTGGCTGGCCT
CTGGTTCATTCCTGTAATCTCAGTCCTTTGAAAGGCTGAGAAAGGAGGATCACTTGAGGCCACAAGTTCAA
GACCATCCTAGACAAGTCAGTTCAAGACCAGACTTCATGTCTACAAAACATCAAAAAATTAGCCAGGCATG
GTGATGCATGCCTGTCATCCCAGCTACTCAGGAGGCTGAGGCAGGAGGATTGCTTGAGCCTGGGAGATTGA
AGTGGCAGTGAGCCATGATTGTGCCATTGCACTCCAGCCTGGGCAATGCATCAAGACTCTGTCTAAACAAT
AATAATAATAATAGTAATAGTAATAATAATAATAATAAAGAAAACGGTTGGGACGCCATTCCTTACTTATT
CAATACACAAAGTTAAAAGCAATTTCTACTTTCTCTATTTTTTTATTACTAAAAAAAGCTGAACCATTCTC
ACAGCCTGAAATGCTTCTCACCTTCCCCTCTTCTATACAAACACTTCTCTGTTGATGATAATGCAGACAGT
CTCTCCTTTAGGAATACTTCACACCAGGTAGTTCCAGATCCCCTTATCTCTGCCTTCCCAGAGCTCCTGGT
GTCTCCCCAGTTCCCTCTGTGTGGTGAAGTACCCCCACCTTGGGTCTCAGCATGACTCGTTCTTTGAAGGT
CTTGTTCACATTTTCCCTTATGGTTCTGTTCCCCTGTGTTGTGTCACAGCACTGGGCAGAGTGGACAACCC
ATTCACACCGATAGAGAGGGCCCCATGGTTCTGGAGATAACCATGTAACTGATCAGAATAGGGCATTGAGG
GCTGGGTGTCAGGCATGGGCTGCACTTGGGTGGGCAGGCCCCCTGGAAAGTCACAGGATTTGGCAAGCTTC
CTAGTAACATCTCTCCCTGGGGTCCTCTTGGAACTTCATGCCCGATGCTGGATGCTGGTTTATTCTCGAGA
GATGGTTCATTCCAATAATCAATGAAACTCATGTCCCAACTAAAGTTCATAAACTCCAGTTTGTAAACTGA
GATAATCTCAGCTGAATGAACTTGCTGACCCTCTGCTTTCCCCCAGCCTCTCAGTGCCCTTGAAATCATGT
CAGTTCAAGCAGCCCCATGAGGCATTACAATGTTTAGTTATTTCAGTGTTTATTAAACCTTGCCCTATGCT
GACCCCAGGTGAATCACAAGCTGGACTTCTGACAAGGACAAGCTATGATATTCTTTTCAATTACAGAAAAA
GTAAGTTAACTGATAGGATTTTTTAAAGATGTTACTAGTTTTGGAAAGGTAATTTGTGCACATGGTAAACA
AGAAGGTATAAGAGGATAATGCTTTTATACTGCTGAGAATAATGTTTTCTCTAATTTTTTTTGGTAGATGC
TTTCATCAGATTAATAAAATTCACTGCTGTTAGTTGTTGAAGGTTTTTTATATCATGAATGGGAGTTGAAT
ATTATCATGTATTATTAATATATTATTATTGAACTAGCAAAGGCTCTTCCTAAAACAATTGAGATGATCTT
ATAATCGTTCTCCTTTAATCTGTTGATGAGATCATTGGTATTTATACTTTTTCTCTGTTAACTATTCTTGA
GTCTCAGGTTTAAATTCAACTTGGTCATGGTGTATCATCTTTGAACACTCCTGTCTCTGGCTTGCTACTAT

TGTGTTCAGCATTTTTGCACTGATGCCGATGAATGAGACTGGCATGTCATCTTCCTTTGCGGTCCTGATTT
TTTTCAGATTTGGATCATGTGGCCCTCATTGAATGAGTTGGGTGTGATGCCTTCTTTTTCATGTATCTGGA
TTGATGGGACACTTTGGAGTCTCTCCAGATGGCCCTCAATGGTCCCTGCCTCCTCATTGTTAGGCTCCTAG
GCAACCCTTTCTCATTTCTGGTAGGCCCAGGAACCTGTGGGTTTTATGTTTGTTTGTTTGTTTGTTTGTTT
GTTTTTTGAGTTGGAGTCCTGCTTTGTCTCCCAGGCTGGGGTTGGAGTGCAATGGCCTGATCTCGGCCCAC
TGCAACCTCCACCTCCTGGGTTCAAGTGATTCTCCTGCCTCAGCCTTCTGTGTAGCTGGGATTACAGGCAT
CCACCACCACTCCTGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTTACAATATAGGCCATTGTGATC
TCTTGGACAGGCTAGTCTCAAATTCCTGACCTCATGATCTGCCTGCCTCAGCCTCCCAAAGTGCTGAGATT
ACAGTTTTGTGCCTCCACACACAGTGAATCTGTGGTTTTTAAAAGCTCCTCATGCATGTGAATTCTGTGAG
CATCCCGGGATGACAGCCACTGTGTGTCCAGCTGTTAAAACTGTGAGAAAGCACCAGCGGGACCCTCTCCA
GCATTTGCTTGCTGTGGTCATGAAAGAGGCTTGTTGGGGAGATGATGCCCTGGTTGACTCCTGAAGGATGG
TTAGGAATGCACCAGATGGAAGCTGGGTTGGACCCAGTCTATGCTAAAGAACAGCTTGTGTGGACACAAGG
AGACACGAACACATCATTTTTGCAGAGCCTGGGGAGTAGCCAATCGCACCATTTGCTTAAAACACCGTGTA
CAGTTGGAGAAGTGGACTGAGACAGGCTGAAGAAGCTAACAGTGGCCAGATGAGAAAGGGTCTTGTGTTAC
TTCCTAGATATACTTAGATTTTATCCTGTGAGTGATAGGAACAGTTGCAGGGACTGAAGCCAAGGAAGCAT
GCTTTAAGATTCCATGTTTTTTGAGATGCTGTCTGGTGGCTGAGTAGGGAATTCCCTGGATAAGTACTGCC
CAGGGTAGGCAAAAGAAGCTAGGAGGTTACTGAAATAAGGAGTATGAGAAATGGTGTAGGTTTTGCTGATG
TTTTGTAACACATCTCATGACAATCTTCATTTCCTTCACCAATTTCCTGTTTCATTAATTCCCTTCCACGT
GCTCTTCTGAAATTTGCCTCATATTCTTTGATTTCTCTTTTACATGTTGGTTTCATCACCTTTTACTTTTT
GCTTTCCTGGAAACACAAATGATTCTGATTGTGACATGTCAGAATTATTTGCAACATTCCCCTTTCTGCTG
AAACATGAGCTCACTGAATACACAATTTAGTAAAGTGTAGGATGCACATGTTGTTTTCATGGTCATAACCA
GCTCTGTAGCATTTTATAACTACACTGGCAGTGTGCTGGGAGGTGTAGAGAGAAATATTTATCTCATGTGT
GGCTGACACAACCTGCCAAGTTGTTTTAGGAGCCTTCTTGGAATCCCAGCAAGAACACCACTGATGCAATT
TGAAATCACAATGTCCTGCTCCATGCCCTGGCTTCATGGCTTAGTCACGTCTGAAGTCTATTTCTAACTAT
CTGTTTCCACATCTATAAAGTATGAGTTAAATCATCCTAATACTACTCATCTTACAAAGTTTTCTTGCTGA
TATTAGGAGAGTTGGGAAAGAACTGTATAAATTATGAAGTGCCATGGAGATGTTGGTGGTTACTTTATCAA
GAAATAGACACTCCAGAATAGAGTAGAAAGAAAACAGTTATGATTAAGTCCTCCTCCTCTTCTTTTTTTTT
AATTTACAAAGAAAGGTTTAATTGAGTCACAGTTCCATATGGTTGGGGAGGCTCAGAAAACTTGCAATCAT
GGCAGTTGGCAAAGTGGAAGAAGGCACCTTCTTCACAAGGTGGCAGGAGAGAGAGAGCTCCTCTTCTTTTT
TGTTGTAAAGTCTACAGAAGTGCATATACTTCAGGGCAAGGGCAGGCAGGGAGAAGAAAGGACATTGCTTC
ACCCCAGTCCTCACTGACAAGTTTGCTTTGGGACTTCATTTTGTCCCAGCATATGGACAGAGCTCTGGCC
ACTACCCATTCAGAAGGCCTGAGCTGCATTGCTAGTTCCCCACTAACTCTGTGTGTCCTTGGGCAAGGCTG
GGCTTATGTCAAAAGATTATGACCCTGGGCTCTCCAGCTACAGAATCTACATATGAATCCTGGCTCTGCTA
GAGCAATTAGTGACGTAACCTTGGATGGGTCAGTTAACCTTCCTGTGGCTTAGTTTGCTCATCTGTAAAAT
AGGGATCATAACAACATCAATACCATGGGTTGTTAGACAGATTGAATCAGTTAATGCAGGGTAAATACTTA
GCATGACACGTATTCACTATCATTTCCTTGAGTAAAAGCTGAGTGTGAGTGGGTGTGAGAATGTGTGAAAC
CCTTTCACTGCAATCTCAGTTAAGAAACCCATCCATAATTTAAAGTTCAGGGCCTAAATGGGTGGTTATCT
TCTCCCAGTTGCATCCTATCCCACCTTTGCTCTTCTCCTGCCCGTAGGAGCTGTTGGTCTTTGATTGGGCT
GGAAGACCTGGTGGACCCTAAGTGATCTATAAGAGAATGAGAATAGAGGACAGGGAATGTCTTCAAAACTC
CTAGAGGGACACAGAGGCTGAGAGGCAGGCAGTCCTGCAGGGGTCTTCTGATTGGGACAAGGAGGACCTTG
GTCTTCATAGGCCAATTCTGGTCAATTTCCCCCATGGACAGATGAGGAAACAGATCCAGGAATATCCAAGG
TCTCACACTTCCCATCTGTCAAGTCTTGTTGATTCTGTTGTATTCATGTCTTTCAAAGAGAGAGAGAGTTT
AAGGAAAGAAAGAAGGATCAACTGTGTCTGATATCACTGGGAGCTTAAGTAAAGGGTTCTTTTACTTCATA
GCATTTTTCCCAATTTGTAATTCAGTATTATTTTTGTCACTGTTTAGTATCTCTTTGTCCTATTAGAGAGA
TAGCTTCATCAGGACAAGGATTTTGATTCTTTAATATTTAGTGCTTGCCACATGCCCTGAACACAGCAGGC
ATACAGACTAACCAACATACAGTGGCATCGAAGTGAGACACCTACCTCCTCCAGTGCCTAGAGTACATGTC
CATGGACCTGTCACTCTCCTCAACACCACCCCTAAGCATGAGGCCCGAAAGCATTGCTAATCCCCTCCTCC
AGCCACCAAAACTTAAAGGCCAGGTGTGGTGGCTCCTATCTGAAATCTCAGAACTTTAGGAGACAGCAGCA
GGAGGATCACTTGAGGCCAGGAATTTGAGACGAGCCTGGGCAACATAGCTAGACACCATCTGTACTAAAAA
TTAGCTGGGCATGGTGGTATACCTGTAGTACCAGCTACTAAGGAGGCTGAGGTAGGAGGATCACTTGAACC
CAGGAGGTGGAAGCTACAGTGAGCTATAACCACAGCACTGAACTCCAGCCTGAGCAACAGAGTGAGACCCT
GCCTCAAAACAATTTCAAAAATAAATAAATAAAAACAAAACTTAGATACCACGTGGTCACCCCAACATGCA
AAATCAAGTTTTCCCCTACTGAGAAGAATGGGGACTTGAGAGCTGAGTTACAGAGAGATAATCTGCCTTTT
TTTTTTTTTTTTTGGTTTACATCCTCAAGATCATGACCTGTGAAATTTGAATCTAATACACAAATCATTCC
AGAGCAATGTTGCTTCTGCCTACCACGAGTAATTCACTTGGCCACTGGAAGTCAGAACAAGCTTCCCAGAA
GAGAGGTACCACTTGGACTACCAATATAAAAGGATGAAAATATCGGAGTGAAGGTGTTCCTTGCATCACTG
AGTCCCTGGACAGCCTGTCCACTCATGCTGATATCTGAGCCTAATGCTTCTCTGAATGTTGAGATTTAACT
TTGATCCAATGAAACCAGACCAAGAAAGAAGAAACGTCTTTCATTGTTGATAAGGACATGATTTTTCTCAC
AATTTTATGATTATTTTTCCTTAGCTGTCCTATAATTATCTGCTTATTTGTCTCTTCTCCATGTGCTTAGG
GTACAAAGTTGACCAAGACCAAGAATAATGTCTGGGAGCACAATACTGACAGCACAGCTTTAAAAACATGA

TGAATGCTTTAATACAGGAAATGAGTAGGGGAGAGGCAAGTGGTGCTTGGGTGTTCTTCCAATGCATAGTA
TCTTCCTTGACACAGTCAGTGCAGCTCTCAGTAGGCAAGTCCCTACATGTTAGAAGATGTTACTTTCTGTG
GAATTAGGTGGCAGAACATGCCTTCAATTATTTTCCTTTGCAGAACAACACCAATTTCATTAGTTAGGACA
GAGTGCTGGCTGCATTTGAATTCCAAGCAACGATTAGTCTATCACTGTTGGTATAGATTCCAACCAGTCAC
ACCACCTCCTGAAGTTTGTTGGGCAGGTAAATCTTCATCTTAGAATAAAAATCATCTTAGCCAAGTAAGTG
TTTTAGAGGAAAGAAGAAAACATAATCGTTTCCATAAGAGTTTTGTTTCTAAAAAAATAAGAAAGGCTCTT
TGTTTAGGTGAGCTAATGAAGTTGTTGATAGTTATCAGATGACACTGGAATCTTTACTTGCCAGAATGTGT
TCTGTGCACCTCTCGGTGTGGCAACATAGAGAGGGAGATCCTCCAGCAATGCCATTGATATGGTCAGAAAC
TGCATCTTTCTTTCTCCCTGCTGAGATGGGGTCCTTTGTTCTAGAAAACCCAGGGGGTGCCACTGGGAGTA
ACCCTTGAGACAGGAACACGAATCTCAACCAATTTCTGGTTGCAGCCTTGAGTCTTACTATTTGCCATAGT
GATGCTTAGCAAGGAATGGCAGGTGCACCAGAGCAGCAGAGGACCTAATATCTCCCTTCCTGTTAACTTTT
TATAATATTTTATTGTGATCAGTATCAGTTGGGAAGCTACTTGCAGTCACTGAGCCTCAGTTTCTACATCT
GTAAACTGGGGATAGTAGCATGGCCCTATTTAATGTGCTCAGCGAAGCCACTGAAAGGAGACAGAAATGTA
CCAGAATTCCCTGGACTTTTATCCTACTTCTCCTGGGGATTGTCACCCACCTACCCGTGTCTGTCCTTTGT
TGCTTTGACGCTGTCACTTCTTTTCTTAGGTACCTCTCTGTAGGGCTCCATTATTCCAGGGATTCCAGAGT
TACAGCACATGCATACCTCCATCCAAGCATGTTTATTTGTCTCCTGCTTCACTAGGCTGTCCCCAAGGAAC
ATGTGGCTCCCGGCACATACCTGGCACAACACTGCACATGACATTCACCCACTTGGCCTTGAATCTGACAA
GGAATCTGGCATGATGTTCACCTGCTGAGGCCAGGTGCCGAGCAGCCCTGGAGGCTTAGGGGCCAGAGGGA
TGGGAAAAGGTGTCTTTCTGGGGTGAGTATCAGTTTCTGCAGGAGTGCTGAACCTGAGAAAGAATAAAGAG
AGAAGGAAGTGAACAAGCACAGCTTAAACATCATCTGTTTCTACTGAGTTTTAACAACTCTGAGATTTTGT
TTGTCATGGAATCCATTTCTCAGGCCAAGCAGACACAGAACTTGGGTGTGAGTGATGATAATGAGCTGATA
TAATTTTCACACCCTCATCACTGAGATCTCTCCCATCAGGAATGGGTCACAGGGCTCACAGGTGGCAGCAA
CTGTTATTACAGGCCTCATCTCTACCAGCTCCTGGCACCTGCTCTCCTCTCATTAGAAAATCCTCCACTTG
TCAAAAAGGAAGCCATTTGCTTTGAATTCCAATTCCACCCTCAAGAGGCTGGGACCACCTCATTGGAGTCC
TTGATGCTGTGTGACCTGCAGTGACCACTGCCCCATTGTTGCTGGCTGAGGTGGTTTGGGTCAACCTGGCC
ATCTGGGCAGCTGTTCTCTTCTCTTCTTTCTCCCCTACTGTTTCCAGACATGCAGTATTTCCAGAGAGAAG
GGGCCACTCTTTGGCAAAGAACCTGTCTAACTTTCTATCTACGGCAGGACTTTTGAAAGCTACAGAGGAAG
AAGCACAAATTGATGCTATTCCACTAAGCCATCAGCTCCATCTCATCCATGCCATGTCTCTTTTTTAGGGG
TCCTCTTGCCAACAGAATCACAGAGGACAAATCTGAAAGTGCAGAGACAGCAGCTGAGGCACAGCCAAGAG
CTCTGGCTGTATTAATGACCTAAGAAGATGGAGTGGTCACCAGAAAGTCAGAGGAAGTGACACACAGGGGC
CCAGCAATCTCAGCCAAGTCAACTCCACCAGCCTTTCTGGTCCCCACTGTGTGTACAGCACCCTGATAGGG
ACCAGAGCCATGAGAGTGAGTAAGACCAGACTATGCCCTTGAGGAGCTCACCTCTGCTAAGGGAAACAGGC
CTGGAAACACACAATGGTGGTAAAGAGGAAAGAAGACAATAGAACTGCATGAAGGGGATGGAAAGTGCCCA
GGGGAGGAAATGGTTACTTCTGTGTGAGGGGGTTGGTGAGGAAAGACTCTAAGAGAAGGCTCTGTCTGGCT
GGGTATGAAAGGATGTGTAGGAGTCTTCTAGGGGGCACAGGCACACTCCAGGCATAGGTAAAGATCTGTAG
GCATGGCTTGTTGGGATGAGTTTCAAGTATTCTGGAATGAGGACAGCCATAGAGACAAGAGGAGAGTTAAT
AGATTTTATGCCAATGGCTCCACTTGAGTTTGTGATAAGAACCCAGAACCCTTGGACTCCCCAGTAACATT
GATTGAGTTGTGTATGATTCTACATAGAATATTAACTCAATGGAGGTCAGTGAGTGGTGTGTGTGTGATTA
TTTGCCAACTGCCGAGGTGGAGAAGCCTCTTCCGACTGCAGGCAGAGCACGGGGGCCCTGCTACTGGCTGC
AGCTCCAGCCCTGCCTCCTTCTCCAGCATATAAACAATCCAACAGCCTCACTGAATCACTGCTGTGCAGGG
CAGGAAAGCTCCACACACACAGCCCAGCAAACAGCAGCA

Figure 9.

GCCAATTAGAAGAAACACAACTACAAGGTCAGGGCATATTATTCAAACAGTAGAGACAATACAGTCAAATA
TTTGGCAGAATTACAAAATATCTCATTGGAAAAGACACGCAAGGGAAATCAACAAAAAGATATGAATCAGA
ATTCATCTGTGTCTCAAGAAAAGGTCATGCGATAAATTAAGTTCTGCTAGTGTTTCTACACTACCGTTAGC
CTCATTACCTTATTTTTTAAGTGTTAATATAGTTTTAGGTATTTTACATACATTTTTATTATTAATTACAA
CCAAAGTGCAACTTGTAATAGCAATTCCTTCACATTTTTTTTTCAAATCTTGCACCTTAAAATCCACCTC
GGGCCTCAGTTGGCCAGCTTTGGTATCTGATACTTGGACTACAGATACCACTAAGGCAAGTAGATAAAATG
TACTCTAGGACCTACAGCCCTTCTGCTAGATCCTGAAGAATGATCATTAAAACAAGCTGGTCTAGCTGGTC
AAGAGCAAAAATAAAATCAAGATGACAGAAAATTGATGCAAAAGTGAAGTAAAATAGCTAGAGAATATGAT
TGCGCCTGTCCCCTTAGCATGGATTCCCATGCTAGCCAATCTAAAATCCTCACTGTTAGAATCCTCCTGTC
AATATGATAGAATGAACAGCAAGCTCAGTGTCAGAAAACCTGTGTTGTTAACTTGGCCCTCTTTCTAGCTG
AATGTGTGTTTTTGGTCAAGTTCTTTGGCATTTCAGAGACTCAGAGTAGTGAAGGAAGTGGATAAGATGAC
CTCTACATTCTCTTGCAAGCTCAAACATCTATGAATCCAGAGAGAAAAACTAGAGCATGAAATTAAGGTTA
TTTTAAAGAAATAACCTTAAAATTATTAGTATTCGAGGATCTCCAATATATTCATGGCACCACTCAAAACT
TTCCTTCTGCTCTATCCCGTCTTGGCTCAAAGTTATCTCCTTAATGAGGTCTGCCCTGACTATCCTACTTA
AAATTGTAAACTTTGCCCACCTGGTACTTCCACTCTCTTTCCCCTGCTCTGTTTTTCACCGTAATACTTTA
CTCTTTTTAACATACAAAATCACTTATTTACTGTGTTGTTATCTATCTGCCTACTCTTACCATCAAATATA
AGTTCTACCTAGGCAGGGATTTTTGTATGTTTTGCTCATGGATATATACGAAGCACTTAGAGTAATATGTG
ACATATACAGGGTACTTGATTAATACTGTTGAGTGAATGAATGAGTTTCCAATACAAATTTAAAATAAAAT
ATTTCCTAACTTAAAATTGTAAAGTCAGATCTAACCAACTGTTCATTGGTCTGCTAGCAGTGTTTCTTGTA
TATGGAAATATATTTTAAATAGATATGTCCTGTGAAATAATACTAAGTGTTCTAAAGAAATAAGTGAGTGA
ACGTTACCTCATTGAACTAACTTGACCTTGCTCCTGGGAGAGAGTTCATTTGAGATTAAACAAGTTCAAAG
TCTATGAATCATAAAACGATAAAAAAAACTAAAAGGGAAATGGTGTTTTTATAAGCTCTGCAATTCAAAAG
CCATTTCGGGTAATATTGTTATTTTTATGTCAGGAATTCCTCAGTGCTGATATCTTAGGGCAAAGGGTTTG
GTTATAAATTAAGAGAATGAGGAAATAGGTACATAGTAGGATTGTTCCAACCAAATATGTGTTGAATGTCA
AAGGAATTTCCCTGAGGAATAATCTTCAGAATAATTTGCTAAGCACAGGAGAAAATTTGGCTTATTACTTT
ATAGCCAGATTTCATTTTTAATTGAAACTTCTTTCAAGCAAATCACTTACTAGTCTATTAACAATAACAAC
ATAAACACAAGTAAACATTCGGAATATAGACATCCAGGTACTAAGCTGATTGCTTTACACTCACTGTCTTA
TTTTACAAGTAAGGAGTTTTAGTTGCAGCAAAAGAAATAAATTTTCCAATGTCAAATGACCAGAACTTAAA
CCCAATCTGTTTGGTGCTAAAGCCAATGTTCTTTACTGCAATGTTGGGTTATCTTGTTTCTAAAACTTAAA
TTTATCAGTAAAAGGCAAAATTTGCTATTATTGAGGACATTAAAATCATATTTTTGTAGACTCTGAGGACA
AATCCAACAAAAAAGTTCCAACTATTTCTTGGCAGGCATCATTGAAATTGGTATATAGCTTCCTTGGGTAT
TGACTTTGAAAAGGAAGTTGGTCACTTTAGATATATAAGTTCAGTCTGTTTGTAAAAACAAAATGAAAACA
AAACAGTTGCCTTATATGCTAAAATTATCCTAATCGTTTTCACCTTTAACAACATATACACACAGAACTTG
AGGAACTTTACACGGCTCATCTTCATATTGTCAGCATCTAGCAAAGTACTTGCCACATAGTGATCAATAAA
AGTTTTAGCCAGCCTGGGCAACATAGTGACAGCCTATCTCTACAAAAAAAAATTAGCCAGGCAAGGTGGCG
CACACCTTTGGTCCCAGCTACTTGAGAGGAGGATGTGGGGAGATCCTTTGAATGCAGGAGGTTGAGGCTGC
AGTGAGCTGTGATTGCGCCACTGTGCTTCAACCTTGGTGACAGAGCAAGACCCTGTCTCACACACACACAA
ATTAAGTAAAGAAAAAAAAAAAGAATCAAAGAAAAAAATAATTCCCCAGCTTAAGTCCATCTTTATTTGTTT
GGATAAGCTATAAAGTGTCAAATAATGCTGTTAATGGACATTTCTCTAGCTCTCCCAAAGGAGGAATTGAG
CACATAGTATGTGCTGTATTTTATATACAGAATAAAAATAGAGACAAGATTTCTACCCTCACAGAACTTAA
ATTCTTCAGGAGAATGACACTGAAGTCCTTAATTGGACTTCTCTCTTCTGTATTATCTTCCTCAAGTGGAG
GTATATGGTGCTTAGTTATGAAAAATACCTCCAGGGCTTTGATCTTCTCAATAACTCTTTGAGGCTGATAT
GAAAACAGTAATTAGAAAAAAACCATGTATCCAACTTATATAGACAGTTGATGACCAAAGCTAGAATCCAG
TTATTTCAGCCTCCCATGTATTTTCTTATTACTTAAGGAGAATCTCTATCTCTACCTCTTTCTCTCTTCTT
CCTCTCTCACTTTTCTTAGAAACATGGGTAAGATTTTCAGAAATATGAGAAACTTATTAATAAATGAAAAA
TACTGGGAATTCTCAATGTTTCTTGTTTTAGCCAGTTAATTTTGGCCTTCATTCAATGTGAGTGTCCCTTA
ATAAGGAGCAAACTCCACTGAGAGATAGATACTAATAACCAGGATTCTGAAAATGCATTCTCATCCCCATC
TCCAAACTTTTATAAAAAATATTATAAAATAATACACTTTTAATATAGGAAATTTCTCAAATACAGAAAAA
ATTAAAGTAAACTCAGACCTAACTTTCATCACTAAAAGATAATCACTCTTGACACTTTGATATCTTTATCT
CTAATATTACACCAATTAATTTGCTTGATATAGTGAATATTATGCTATTATAATTTTCCCCTGCCTTTTGT
CCTTGCATATTAGCATAGGTATTTTCTGAGGTTATCACAAACTCTGTAAGCACGTTTTATATTACTACTTT
TTTAAAGAGGATGTATAATAATTAATTCATCCATATATATGTGGTTAAGTATTCAGGTCACTGCTCATTTT
TCACTGTTATAAAATAAAGCAGCAATGAATACCTTTGGCTGATATTTTTTTCTGTACTTGGAATTATTTCT
TTAAGATAGATTTCCTAAAATTGAATTACTGAGTCAAACAGACTTAAGTTTTCCTTATGTATGTTTCCTTA
TTCATTTGAATAATTTTCAACTCCTACTTGTTTATTTAACTCTTGTGAGCATGTGATAGTCTCATTTTTCA
AAATATCTTTGCTGTTGTAATTTGCATTTCTTTGTAGTTAGCATGAATATTTCAGTATGTTTTCTTCTGTG
TACCAGTATACTACATACTTTTTTATATGAATTGCCTATTTGCATCTTTTGCTCGGTTCTATTAGACCTTT

GAATTTTTTCTTATCCATTTATATAAGCTCTTTATATATTAAGAATATTAACCTATTGTGATATTTGCAAT
AAATAGCTATATGGTTTGTTGTTGGTTTTAAATGTGAATTTATTCAATTTTCTTCATAATTTTGTTGTTTT
TATAGATTTCTTTAAAAGTAATAAAATTATTGCCATATTATTATTTATTTTCAATGTCCATTACTAGCCCT
TTCCAGTCATCACTTTTACTCTCATGTTCTTAATTTTTATTCATATCTTGGCTCCATTGACCATCTTTATT
AGGATATTTGGGAAATAATACAATTTATACTAAACACACATCAAATCTTACCTTATTTTCTTCTTACAAA
ACCCTAGGAATATGCTGTTTTTGTCTTTATTTGAATGACACAGAAATCAAGGTTTTTGAGCAGTGGAGTAT
TTCTTCAAATGACACAGAAATCAAGTCTTTTGAGCAGTGGAGTATTTCTTCAAAGCCACACAGCTAGTAAG
TCATGAAGCTGGAATTCCAAGAGTTGCCACTTCATTTTCTTCTTTCCCTTTATCTTACTCAGTTGTCTTCT
CTCCTCTTAATTTTGTCATTCATTTAAAAACATTTCTTGTGCTATTATGGTAGATTTATTTTAATAGGGGG
CAGTGACTTACTCAGAGAGATGATTCTCTAATGGAGTTTTAAAGATCTTAGAAGTTGATAGAGGAGGCTGG
GCGTGGTGGCTCATGCCTATAATCCCAGCACTTTGGGAGGCTGAGGTGGGTGGATCACTTGAGGCCAGAAG
TTCAAGATCAGCCTAGTCAATGTGGTGAAACCCCATTTCTACTAAAATTACAAAAATTAGCCTGGTGTGCT
GGTGCACTCCCATAATCCCAGCTACTCAGGAGGCTGAGGCATGAGAATTGCTGGAACCCAGGAGGCGGAGG
TTGCAGTGCGCCTGGATTATTAACACTACACTCCAGCCTGGGAGACAGAGTAATACTCCATCTCAATAAAA
AGAAGTTGATAAGGGAGATAGTTCATGGCAACGGATCTTTGAAGGCACGCTAATGATAACTTAGGCATTTA
GCCTACTAGTGTAATTTCCATAAATCTGCCTCTGATGTCATACTCTCAGCACCTAATATTTTCTACAAACA
TTTATTGAAACTTTATTTTGTATAAGTCTCTGTCCAGTTTGAATATTTAAAAAATTCATAATCATATGAAA
CATTAATAATAAATACAAAATGAGAGATGCCGATACTGAAAAGTAGGATTGCGGAGTGGTAGAAAATATTT
CTGGCTGTAGTAGATGGGGAAGTGTTCAAAGAGGAGTATAATTCAGGTTTCCATTTGCCATCGACTTATCA
CATGGCTAACTCACTAAGCGACTTAATTAAAATTAAATTAATTTATCATCATCTGATCACCATTTCACACA
ACTCATGTCTGTTGCTGTATTGGCTAAATGATGGCAAGACAAACGACCTCTGAAAATGATCCTATTGACCT
TCGGAATCTGGATTTTTTTTTCAATGCAGGTGTCCATAGAAGCAATCTGATGTAATCCAACATGAGTTCAA
GCACAGTCATTTAATATCCCCTATCAAGTACAGTCATTTAATATCCCCTATCATCACATGTCCTTCATACA
TAAAAATCATTACATGTGAAAGGGTGGAGAGTGTGTGGATCCCTTATTATTGTGTTATTGTAACACAATAA
CAATATTGTGTTATTATTGTTAACACAAGTGAGTCATATGTCTTGCTCTTTGGACTGAGTGGAAACTTGTA
TTCTTTCTCTGCCTCAGGTCAATTAACTTCATTGAGGTGAGTTGCATTCCTTCTTTAAGCATGTTGAACCT
TCAATCTGGACTCAGATGGGCTAAATAGAGGAGCTAGGAAAAATACAGAAAATAAATTATTAGAGAGATCA
GAGAAAGATACATAAGATTTACAATAAAAGAATTATGAGAAAAACATCCAAAAGAATTAAAAACCATAGGA
GAAGGAAAAATAGGTGAACAGCTTTTTAATTTCTATAAATGTGTTGTTAATACTCATAATAAAGGACTCAG
AGCTGGGATATGAGAATAATAGGTCAAACGTATATGGATACATAGATGTGACTACATACATGAGTTGCAAA
GAATGCTAAGGAGGGCAAAAAGAGATTGAGAAGAGGGCATTATTACTAATATATAGCAATGTTGAATGTTT
AGGGTGTTTCAGGCACTGTACAAATCTTTTAAATACACAAATCACTTAATCTTGCCATAACATTAGAAGAT
ACTATCTACTCTTTACCAAAAAGGTAACTGTGGGATAGCAGAGTTAAGTAACCCGTTCAAACCTATGCATA
ATAATCAGCAGAGATGGTCCTTATCTAAGTTTTTCTGCCTTTGAAGTCCAAATAGTTTAATGCAGCCAGGT
ACTAAAGAAGAAAACTTTTGTAAATTAGTTTAGTTTAATGATTTACATGTGGAAAAGCACAGAGTGAAAAG
CACATATCATCTGAGAAGCCCAGTGAGTTTGGCTGAAATGGAGTGAACATGTACATGTTGAGGGTGAGGAA
GATGATTAGAGAGAAGTGATGTGTTGTGGTTCTTAAAAGCTAAGAGGAAACTGTTAGATATGATATAGTCT
GTGGCAGGGAGCCATTGCAGGTTTTCCAAGATGGACTCATAAGGAGAAAACACTTCGTGGGACTGGAAAAG
GCAGTGAAGTGGTGTGTCCTATGATAATGTACTACAGTGTAGGATAGTGGTTAAGAGTACAGTTATGAGAG
AGGGACTACTGGTTAGCACCTTACCTGCTGTGTGACTGGGCAAATAATGCAAACCTCAGTGTCTTTTATTG
TAATATGGGAGTAACAAAAATAGTAACTACTTCATAGGATTCTTGTAAAGATTAAATGACTTAATTTCTTT
GAAGTGCTTGGCAGTTCCTGATAAATGACCAGTAGTTAATAAATGTTAGTTGTTATTATTATCATTATATA
TTATTACTCCCATAGATACATATAGAACAGACTGCAGCAGAGAGGCAAATCTTTAATGTTGTCAGAGTATA
GACAAGTTGGTGAAATGGCTACATGAGAGCGGAGGACAAGAAGGTGCAGATTGTGGCAGTCACTTCAAATG
GAAATATCACCGCTTGAATGAAGGTATATGAGTGTCAACTTGCAAGGGGACCAGGTAGGTTTCATCAGAAA
TTAAGGAAGCTTAAGGAGAACAGCCAAGTTCAGCTTGACAGAAGTGGTGGTGGCACAAATGCAAGACTGGT
GTCTTTCAAGAAACCAAGGACTGTTGAAAGTAGCAAGAGCTAGTTTGTTTTAGGTCCATCATGTTTTATAT
TCACACTTTCATGTCAGTGGAGCAAAGAAATGGAATACAATATAATAGAATGGTAGAATCTTATTTTTAAA
ATCTGTGTTATTCTGATCTTTAACTTACTTATATCTTTGATAGAGATCTTTACCTGATGCTCAAGATTGTA
GAAATAGTATAATCAACATAACAGTATAGCACTGTATTTATATCCTGCACTGTTTAGGGAGGGTTTAAGGC
CATTCAAAAGGATACATAAAATACAACAAGATTACATAAATGAAAGGTGAGATAAAGCAACAAAGCAAAAC
AAAAGTGAAAACAGAGATCATAGGCACAAATAAGATTAAAAACGCATGTAATGAAGATGAAAGCTTTTACA
TTTACCCCAGATGGACCACAGGGTTGTTGTTAAGCCTTTAAACAGTGAACAATGCTGTACACTTGCATATG
CAATTAGAACATGTGGAAAAAAATAGTGGCCTGTTAGAAGCCTAATTAACAATTTGTGAAAAAAAAAAAAAA
AAAAAAAAAAAGAGGCCGAGCTGTAGCTCACGCCTGTAATCCCTGCACTTTGGGAGGCCGAGGCGGGCGGA
TCACGAGGTCAGGAGATCAAGACCATCCTGGCTAACACAGTGAAACCCAGTCTCTACGAAAAATACAAAAA
ATTAGCCGGGCGTGGTGGCGGGAGCCTGTAGTCCCAGCTACCTGGGAGGCTGAGGCAGGAGAATGGTGTGA
ACCCGGGAGGCGGAGCTTGCAGTGAGCCGAGATCCTGCCACTGCACTCCAGCCTGGGCGACAAAGCAAGAC
TCCGTCTCAAAAAAGAAAAAAGAAAGAAAAACAAAAGAAAACTTCATTGTATTGTAAGGCCAAGAACAAAA
TATATCAAGATAAGGAAAATTTGTAGTCAAGAATAGAAAAAAATTATGGCTTTGAAGTATGAGTTATTTAA

AGAAAGTGGAAACATCCTCAGACTATGCAGTAAAAAACAAAGTGATTTTCTTCTTCTAAACTTATGCAATA
AACTGATAGGTAATATGTGAAAGTCATAGAATGTAGACTAGAGGATACAACAAACCTATTTCCTCTATGTT
CATAAGAAGTAAGAAAAGCTCTGATGTGAGTTAGCATTGCTTTACAATTTTGAATTGTGCAGATTGCACGT
ACTTTTCCTCAGTTTGAAGTAAATAGTGGACAGGAAAAAATATTAAATGTTGGCAGTAAATATGGAAGGAA
ATTACAACTAATGTAATATGCTAAAACATGCTATGTTTATTTTACTAATTTGAATTAAAATGTAAGAATTT
AAAATGCCCTGGAAAAACACGGGCATTGATCTGACGTCTGAAGTTTTAAAATATTACACACTTTGAAATAG
CATTTGTACCTTGAAATACCTGTCTCTATATATTTTTTAAAACTTCCTTTTTCTTTCATTCCATTTATCAT
CAAATAAAGGATGAACAGATGTAACTCAGAAACTGTCAAGCATGCTGAAGAAAGACCACTGCAGAAAAATT
TCTCCTAGCCTTTTCAAAGGTGTTAGGAAGCAGAAAGGTGATACAGAATTGGAGAGGTCGGAGTTTTTGTA
TTAACTGTATTAAATGCGAATCCCGAGAAAATTTCCCTTAACTACGTCCTGTAGTTATATGGATATGAAGA
CTTATGTGAACTTTGAAAGACGTGTCTACATAAGTTGAAATGTCCCCAATGATTCAGCTGATGCGCGTTTC
TCTACTTGCCCTTTCTAGAGAGGTGCAACGGAAGCCAGAACATTCCTCCTGGAAATTCAACCTGTTTCGCA
GTTTCTCGAGGAATCAGCATTCAGTCAATCCGGGCCGGGAGCAGTCATCTGTGGTGAGGCTGATTGGCTGG
GCAGGAACAGCGCCGGGGCGTGGGCTGAGCACAGCCGCTTCGCTCTCTTTGCCACAGGAAGCCTGAGCTCA
TTCGAGTAGCGGCTCTTCCAAGCTCAAAGAAGCAGAGGCCGCTGTTCGTTTCCTTTAGGTCTTTCCACTAA
AGTCGGAGTATCTTCTTCCAAAATTTCACGTCTTGGTGGCCGTTCCAAGGAGCGCGAGGTAGGGGCACGCA
AAGCTGGGAGCTACTATGGGACAGTTCCCAAGTGTCAGGCTTTCAGATTTCCTGAACTTGGTCTTCACGGG
AGAAGGGCTTCTTGAGGCGTGGATAGTGTGAAGTCCTCTGGCAAGTCCATGGGGACCAAGTGGGGTTAGAT
CTAGACTCAGGAGCTCCTGGAGCAGCGCCCAAACCGTAGTGGCACTGGACCATGTTGCCCGGAGCGCGCAC
AGCCCGCGCGGTGCGGGGACCTGCTCTCTGAGCCCGCGGGCGGTGGGTGGAGGAAGCATCGTCCGCGGCG
ACTGGAACCGGGAGGGAGAATCGCACTGGCGGCGGGCAAAGTCCAGAACGCGCTGCCAGACCCCCAACTCT
GCCTTCGTGGAGATGCTGGAGACCCCGCGCACAGGAAAGCCCCTGCAGTGCCCATCGCGGCCAGAGCAGCT
GGGGCATCAACGGCGGGCGCTCCCTCTTACTGCTCTCTGGCTTCGACGGGGGACTAGAGGTTAGTCTCACC
TCCAGCGCGCCTGAGGCTCATGCATTTGGCTAATGAGCTGCGGTTTCTCTTCAGGTCGGAATGGATCTTGA
AGGGGACCGCAATGGAGGAGCAAAGAAGAAGAACTTTTTTAAACTGAACAATAAAAGGTAACTAGCTTGTT
TCATTTTCATAGTTTACATAGTTGCGAGATTTGAGTAATTTATTTCTAGCCTCCAGCTCTGAAATAAATGA
CATGTTGTTGTTTTTAATTATTTTTAAGAAACGCAAGCTAGCCTTTG

Untitled.ST25.txt
SEQUENCE LISTING

<110> THE UNIVERSITY OF SYDNEY

<120> THE UNIVERSITY OF SYDNEY

<130> P41173

<160> 5

<170> PatentIn version 3.1

<210> 1
<211> 12983
<212> DNA
<213> Homo sapiens

<400> 1
ctggttcatc tcattgggac tggttggaca agagggtgca gcccacggag ggtgagccaa
60

agcagggtgg ggcgtcgcct cacctgggaa gcacaagggg tcgtggaatt ttctccccta
120

cccaaggaaa gccataaggg actgagcctg aggaactgtg cactctggcc cagatactgc
180

acttttccca tggtctttgc aacccgcaga ccaggagatt ccctccggtg cctatgccac
240

cagggccctg ggtttcaagc acaaaactgg gcagccattt gggcagacac cgaactagct
300

gcaggagttt tttttttttt tttccatacc ccattggcac ctggaacgcc agtgagacag
360

aaccgttcac tcccctggaa agggggctga aaccagggat ccaagtggtc tggctcggtg
420

ggccccaccc ccatggagcc cagcaaacaa agattcactt ggcttgaaat tcttgctgcc
480

agcacagcag cagtctgaga ttgacctggg accctcgaac ttggttgggt gctgtggggg
540

ggcatcttcc attgctgagg cttgagtagg tggttttacc ttcgcggtgt aaacaaagct
600

gctgggaagt ttgaactggg tggagctcac cacagctcag taaggccact gtggccagac
660

tgcctctctg gatttctcct ctctgggaag gatatctctg aaaaaaaggc agcagcccca

720

gtcagggact tatagatgaa acccccatct ccctgggaca gagcccctcg gggaagaggt
780

ggcttccacc attgtggaag actgtgtggc aattcctcac ggatttagaa ctagagatac
840

catttgaccc agcaatccca ttactgggtg tatacccata ggattataaa tcattctact
900

ataaagacac atgcacactt atgtttattg taacactatt tacaatagca atgacctgga
960

accaatccaa aagcccatca atgatagact gaataaagaa aatgtggcac atatacactg
1020

tggaatacta tgcagccata aaaaaggatg agttcatgtc ctttgcagag acatggatga
1080

agctggaaac catcattctc agcaaactag cacaataaca gaaaaccaaa cactgcatgt
1140

tgtcactcat aagtgggagt taaacaatga gaacacatgg acacagggag gggaacgtca
1200

cacactgggg catgtcgggg agtgggggcc tacgggaggg atagcattag cagaaatacc
1260

taatgtaggt gacgggttga tgggtgcagc aaaccaccat ggcacatata cacctatgta
1320

ataaaactgc acgttctgca catgtacccc agaacttaaa gtataattaa taataataat
1380

aatttctggg catgtaagta gctgtctttc aggttctact ttgatacata ttctgagaga
1440

attaaacctg tcaaagaaac cttgactttc aatggcaggc actggaattg accctaataa
1500

tgtgttttgg ggtaagccta ctcatattct caacctgtct gcagtagtcg ttagaatctg
1560

aacttcctga agttcatgtg caaagttgag ttaattgttt aatattcaac aaggattatg
1620

ccagtaagat ggtaggaaaa tattagatat gtgtcatcac tgctggtatt atttaaactg
1680

caacatattt tagctggctg ctgatctcag ccaccatgcc tgcattttat ctctgtctcg
1740

tggtctgcaa ccttggaagc tttgaactta gctcatagaa tcctgggcat caagaacatg 1800

tggttctaat ggctagatag ggaatgagag taaaaggatt ttgcccacgg tcacgtgagt 1860

aaacaacaga tttggagggg tctggactac tgtgatgact tcattctgac aatatgttcc 1920

agttgtcctt tcatttcctc ctaatcacat gtctggtctg atctggctgt ttcccacctt 1980

ccaattcctg ccttctccaa tgctcccttc cgtaggtcac tctgtggctc agagaccctg 2040

cttagcaagc gcccaacctt tcaattattt gttcagtaaa acttgaactc atgtctcccc 2100

ttcttgataa aaagaaaata cgttatgtaa tgtcgggtta ctctataact cttgtcctgt 2160

ctctcggcaa ctactgaact aactgttttc atattgagca aacgtttatg gaaggactgc 2220

caagagtcag gtactaggct tggtaatatt ccccgttctc tctagtcaaa gccaacacca 2280

gccagacttg cagatctagg tcccaagccc actgcagatc acaggccagg gtctggtctc 2340

ctctgagctc ctttgggagg gaaagacaga attattaaca cccattttgt agattaggca 2400

actgaggctg aggaagttta aataactcag acagggcctg cacgtcagtc atattccaag 2460

gatccctact cactgtcttc tctctacaga acgagatgtc tctggagtcc atagaaagcc 2520

caggagcctg gctgggcacg gtggctcctg cctgtaatcc cagcactttg ggaggccgag 2580

gcaggcagat cacctgagct caggagttca agaccagcct gggcaacatg gcaaaacccc 2640

atctctacta aaaatacaaa aaattagctg ggcgtggtgg tgcatgcctc taatcccagc 2700

tacttgggag gctgaggcac aagaattgct tgagcccagg aggcagcagt tgcagtgagc 2760

tgagattgtg ccagtgcact ccagcctggg caacagagca agattccatt tcaaaaacaa 2820

aaacaaacac aaacaaacaa acaaaaatag aaagcccagg gaccacctgc gtcaggttcc 2880

cagccacacc tttttcttgt cctcctctgt ctctggcatc ttctcacagg ttcctaattg 2940

tttgtggttg cacaaattca aaatcccaga aaaattacca cttcacaccc actcagatgg 3000

ctattttttt tttgaaggaa gataacaagt gttgacaaga acatggagaa attggaattc 3060

tcacccattg ctggtgagaa tgtaatacgg tgctgctgct atggaaaaca gcttggagtt 3120

tcctcaaaaa gttcaacaga atttcaatgt gacccagcaa ttcccctcta agttatagat 3180

ctgagaggat taaaaacagt tactaaaata cacggactca catatttcta acagtccaat 3240

tcacaagggc caaaaggtgc taatagccca catgtccatc gatggatgga taaataaatt 3300

gtggtctatc catacaatgg aatattattc ggccataaat ggaatgaagt actgacgcat 3360

gctacagaat ggatgaaccg caaaaaaaat ggatgaacac atgctacaga atggatagcc 3420

tcactttact atgaagtgaa ggccagaaac gaagtccata tattgcatca tacaaaatat 3480

ccagaagagg gaagcccaca gagacagaat gtgcaatggt ggatgccagg gtctggggag 3540

aggggagagt ggggagaaac tgctcaactg gtacaggctt tattttggaa tgatgggaac 3600

attttgcaac tagatagagg tagtgattgc agaacacaga atgtactgaa ttccactgat 3660

ttttttcacc ttaaaatggt taattttcag tcctgagatt ggataatcat aaaaaaatgg 3720

ttaattttat gttatgtgaa tttcatccct atacatattt taaacctcag aaatatacac 3780

tagcaggcat ggaacaggtc actgtggtgc ctgccaagcc cggtgatgtt atctggggtc

3840

cccggccagc cttaagcctc ttgctgaccg gtggagggca gaacctttgc cctaaaagta
3900

taatatccac atgctggcat gattcctggc cagatggctt ctttattagc agtaattgaa
3960

actgcctcga tacagacact gtaccttgca accaaaaaat gactcaacaa tgataataag
4020

ggttaagctg ggcctttctc tctttgccag ttaaattata tttattatag cttgacatga
4080

aaaacaaagc aactccaaca ggtatcacaa gggcaaagga catgaacatt ttatcaaaga
4140

agaaatgcag ctgtcaaaaa tacagaaata ttcaaccttg ttcataataa agtggctggg
4200

ctcagtggtt catgcctgta atcccagtgc tttgcaaggc tgagacagga ggatcatttg
4260

aagccagaag ttcaagacca tcctaggcaa gtcagttcaa taccagactt catgtctaca
4320

aaacatcaaa aaattagcca ggcatggtga tgcatgcctg ttgtcccagc tactcaggag
4380

gctgaggcag gagaattgct tgagcctggg aggctgcggt ggcggtgagc catgattgtg
4440

ccattgtact ccagcctggg caatgcagca agactgtcta aataacaaaa ataatagtaa
4500

agaaaaggat tgggatgcca tttacttgcg tattcaatac acagagttaa aagtaatttc
4560

tacgttttct atttttttat tactaaaaaa agctggacca ttctcacagc ctgaaatgct
4620

tctcactttc ccttcttctg tccaaacact tctctatgat aatgcaaaca gtcactcctt
4680

taggaagact tcaccccagg tagttccaga tccccttatc tctgccttcc cagaactcct
4740

ggtgtctctc cagttccctc cgtgtggtga agtaccctac ctagggtttc agtatggctc
4800

tgtctgcaaa ggtcttgttc acaccttccc ttatggttct gttgccctgt gttgtgtcat
4860

agcacagggc acagtggaga acccattcac actgatagag agggccccat ggtcctggag
4920

ataaccatgt aaccgatcag aataaggcat tgagggctgg gtgtcaggcg tgggctgcac
4980

ttgggtgggc aggtcccctg gaaagtcact gggtttggca agcttcctag taacatgtct
5040

ctctggggtc ccccttggaa cttcatgcaa aaatgctggt tgctggttta ttctagagag
5100

atggttcatt cctttcattt gattatcaaa gaaactcatg tcccaattaa aggtcataaa
5160

gcccagtttg taaactgaga tgatctcagc tgaatgaact tgctgaccct ctgctttcct
5220

ccagcctctc ggtgcccttg aaatcatgtc ggttcaagca gcctcatgag gcattacaaa
5280

gtttaattat ttcagtgatt attaaacctt gtcctgtgtt gaccccaggt gaatcacaag
5340

ctgaacttct gacaagaaca agctatcata ttcttttcaa ttacagaaaa aagtaagtta
5400

attgatagga tttttttgt ttaaaaaaaa tgttactagt tttgaaaagg taatatgtgc
5460

acatggtaaa cactaagaag gtataagagc ataatgcttt tatactacta agaataatgt
5520

tttctctaag tttttttgg tagatgcttt catcagatta agaaaattcc ctgctattag
5580

ttgttgaagg tttttatatc ataaatgaaa gttgaatatt attatcatat attattaata
5640

tattgttatt gaactatcaa agccttttcc taaaaccatt gagatgatct tataaccatt
5700

ctcctttaac ctgttgacga gatcattggt atttatacta tttctctgtt aaccattctt
5760

gagtctcagg tttaaattca acttggtcat ggtgtgtcat ctttgatcat tgctgtctgt
5820

ggcttgctac tgttttgttt aggatttttg cactgatgct catcaatgag actggcatgc
5880

catcttcctt tgcagtcctg atttttttct gatttggatc atgtggttat ggccctcatg 5940

gaatgagttg ggcatgatgc ctttttttca tgtctctgga ttgatgggac actttggatt 6000

ctctccagat ggccctcaat ggtccctgcc tcctcattgt taggcccctg ggcaagccct 6060

tctcatttct ggtaggccca ggaacctgtg ggggttttgt ttgtttgttt gtttcttgag 6120

tcggagtctc actctgtcac ccaggctgga gttggagtgc aatggcccga tcttggctca 6180

ctgcaacctc cacctcccag attcaagcaa ttctcctgcc tcagcctcct gagtagctgg 6240

aattacaggc acccaccgac acaccctgct aatttttgta tttttagtac agatggggtt 6300

tcacaatatt ggccaagctg gtctcgaact cctgatctca tgatctgccc ggcttggcct 6360

cccaaagtgt tgagattaca agcatgagcc accacaccca gtgaacctgt ggtttttaga 6420

agctccccat gcatgtgaat gctgtgagca tcccaggatg acagccactg tgtgttcagc 6480

tgttggaact gtgagaaagc accagtggga ccttctccag cacctgcctg ctgagttcat 6540

ggaagaggct tgttggggag atgatgccct ggctgactcc tgaaggatgg ttaggaatgc 6600

accagatgga agctgggttg gacccactct atgctgaaga acagcttgtg tggacacaag 6660

gagacacgga tatgtcattt ttgtagagcc tgaggagtgt ccaatcacac catttgctta 6720

aaacatcatg cacacttgga aaagtggact gagaccgaat gaagaagcta acagtggcca 6780

gatcagaaag ggtcttgtgt tacttcctag agatacttag attttatcct gtgggtgata 6840

ggagcagttg gagggactga agacaaggaa agaaacatgt ttcaagatct atgtttttca 6900

agacgctttt ctggtggctg agtagggaat tccctggata agtcctgccc agggtcaggc

6960

aaaacaagtt aggggggttac tgaaataagg agtatgagaa atggtgtagg ttgtgctgac
7020

gttttgtaac acatctcatg atgatcttca tttccttcac taatttcctg tttcattaat
7080

tcccttccac gtgctcttct gaaatttgcc tcacattctc tgatttctct tttacctgtt
7140

ggtttcatca ccttttactt tttgctttcc tggaaacaca aatgattctg attgtgacat
7200

gtcagaatta tttgcaacat ttgcctttct gctgaaacca tgagttcact gaatacacaa
7260

tttagtaaag tgtaggatgc acatgtcgtt ttcgtggtca caaccagctc tgtagcattt
7320

tataactaca ctggcagtgt gctgggaggt gtagagagaa atatttatca catgtgtggc
7380

tgacacaacc tgccaagtta ttttaggagc ctccttggaa tcccagcaag aatgctaccg
7440

gcacaatttg taatcacagc atcctgctcc atgccttggc ttcatggcat agtcacttct
7500

gcaagtctct ttccagctgt ctgttcccat gtctataaag tatgagttaa atcatcctaa
7560

cactactcat cttacaaagt tttcttgctg atgttaagag agttgggaaa gaactgtata
7620

aactgtgaag tgccatggag atgttagtgg ttactttatc aagaaataga cactctagaa
7680

tggagtagaa agccaacagt tatgattgag tcctcctcct cttcttcttt ttattaattt
7740

ataaagaaaa gaggtttaat tgactcacag ttccatatgg ctggggaggc ctcgggaaac
7800

tctcagtcat agcaggaggc aaaggggaag aaggcacctt cttcacaagg cggcaggaga
7860

gagagagctc ctgttctttt ttgtcataaa gtctacagaa gtgcttatac ttcaggacaa
7920

gggcaggcag agagaaggaa ggacattgct tcaccccagc cctcactgac gagtttgcta
7980

ggggacctca ctttgtccca gagtagggca gaactctggc cactacccat tcagaaggcc 8040

tgggctgcac tgctagttcc tcactaactc tgtgtggcct tgggcaaggt tgggcctgtg 8100

ttaacagatt atgaccctgg gctctcaagc tagaggatct aaatttgaat cctggctctg 8160

ctaaagcaat tagtgatgta aactttaatg ggtcagttaa ccttcctgtg gcttagtttg 8220

ctcatctgta aaatagggat cataacagta tcaataccac atgattgttg gacagattga 8280

atcagttaat gcaggggaag tacttagcat gacacgtatt cactatcatt tcctggagta 8340

agagctgtgt gtgagtgggt gtgagcatgt gtgaaacctt ttctctgcaa tctcagttaa 8400

gaaaccaatc cagaatttaa agttcagggc ctaaatgggt ggttatcttc tcccagttcc 8460

atcctatccc acctttgctc ttcctcccgc ccacaggagc tgttggtcct tgattgggct 8520

ggaagacctg gtggacccta agtgatctat aagaggagaa tagagaacag ggaatgtctt 8580

caaaaatcta gagggacaca gaggctgaga ggcaggcagt cctgcagggt cttctgattg 8640

ggacaaggag aaccttggtc ttcacaggcc aattctggtc agtttccccc atggacagat 8700

gaggaaacag gcccaggaat atccaaggtc tcacacttcc catctgtcaa gtcttgttga 8760

ttctgttgta ttcatgtctc tcaaagggag atagagttta gggaagaaag aaggatcaac 8820

tgtgtctgat accactggga gcttaagtaa agggttcttt tacttcatag catttatccc 8880

aatttgtaat tcagtattat ttgtgtggct gtttggtgtc tctttctcct atatgagtgc 8940

tagcttcata agggcaagga ttttgattct ttaatattta gtgcttgcca catgccctga 9000

acacagcagg catacaggct aaccaacata cagtggcatg aaagtcatga aagtgagaca 9060

cctacctcct ccagtgccaa gagagcataa ccatgcacct gtcactctcc tcaacaccac 9120

ccccaagcat gaggcccaaa agcattagct aatcccctcc tccagccact aaaacttaaa 9180

ggccaggtgt ggtggctccc atctgaaatc ccagaacttc aggagacagc agcaggagga 9240

tcacttgagg ccaggagttt gagatcagcc tgggcaacat agctaggtcc catctgtact 9300

aaaaattagc tgggcgttgt tgcatgcctg tagtcccagc tactaaggag gctgaggtgg 9360

gaggatcact tgagcccagg aggtggaaac aacagtaagc tataatcaca gcactgaact 9420

ctagcctggg caacagagtg acaccctgcc tcaaaacaat tttaaaaata aataagagca 9480

aaacttagat accacgtggt caccccaaca tgcaaaatca agttttcccc tactgagaag 9540

aatggggact tgacagctga gttacagaga gataatcttc ttcttctttt tttttttttg 9600

gtttacatcc tcaagatcat gacttgtgaa atttgaatcg aatacacatg taattccaga 9660

gcaatgttgc ctccgcatac catcagcaat tcacttggct actggaagtc aggataagct 9720

tcccagaaga gaggtaccac ttgggctacc aatataaaag gatgaaaata tcagagtgat 9780

ggtgttcttt acaacgttga gtccctggac agcctgtcca ctgatgctga tatctgagcc 9840

taatgcttct ctgaatgttg agattgaact ttgatccaat gaaactagaa cgagaaagaa 9900

gataagtctt tcattgttga taaggacatt atgtttctca tacttgtatg attatttttc 9960

cttagctgta ctataattat ctgcttattt gtctctgctc tatgtgctta gggtacaaag 10020

ttgaccaaga ccaactttgg ttggaagcat agtactaaga gcacagtact gagagcacag

10080

tattgagagc acagctttaa aaaacatgat gaaggcttta atacaggaaa tgagcagggg
10140

agaggcatgt ggtggttgga tgtatcttcc ttgacacagt cagtgcagct ctcagtagtc
10200

aagtccctac atgttagaag atgttacctt ctgtggaatt aagtggcaga acttgccttc
10260

aattattttc ctttgcagaa caacaccaac tgcattagtt aggacacagt gctggctgca
10320

tttaagtccc aagcgatgat tagtctctca ctgttggtat agattcaaac caatcagacc
10380

acctcctaaa gtttgtaggg caggtaaatc ctcatcttag aataaaaatc atcttaccaa
10440

gtatgtgttt tagaggcaag aagaaaacat atttgtttct gtaagagttt tgtttaaaaa
10500

aaatataaga aaggctctcg gtttaggtga ggtaatgaag ttgttgatag ttatcagatg
10560

acactggaat ctttacttct ctgaacgtgt tctgtgcatc tctcagtgtg ggaacataga
10620

gagggagatc ctccagcaat gccactgata tggtcagaaa ctgcatcttt ctttctccct
10680

gctgagatga gatggagtcc tttgttctag aagacccatg gtggtgccgc tgggagtaac
10740

ccttgagaca ggaacacaaa tcccaaccaa tttgtggttg cagccttgag tctcactatt
10800

tcccatagtg atgcgtagca gggaatggca ggtgcaccag agcaggagag gacctaatat
10860

ctcccttcct gttagctttt tataaagttt tattgtgatc agtagcagtt gggaagctac
10920

ttgcagtcac tgagcctcag tttctacatc tgtaaactgg ggatagtagc atggccccta
10980

cttaatgtgc tcagcaaagc cactgaaagg agacagaaat gtatctaaat taccctggac
11040

ttttatccta cctctcttgg ggattgtcac caccttccca tgtttgtcct ttttggtttg
11100

atgcttgctg tcacttcttt ccttaggtgc ctctctgtac ggctctttta tcccagggat 11160

tccagagtta cagcacatgc ataccaccat ccaagcatgt ttatttgtct cctgcttcac 11220

taggctgtcc ccaaggaaca tgtggctccc ggcacacacc tggcacaaca ctgcacatga 11280

cattcaccca cttggccttg aatctgacaa ggaatctggc atgatgttca cccactcagg 11340

ccaggtgccg agcagccctg gaggcttagg ggccagaggg atgggaaaag gtgtctttct 11400

ggggtgagta tcagtttctg caggagggct gaatgtgaga aagaataaag agagaaggaa 11460

gcgaacaagc acagcttaaa catcgcctat ttctattgag ttttaagaac gctgtgattt 11520

tgtttgtcat gcaatccatt catcaggcca ggcagacaca gaacttgggt gtgagtgacg 11580

ataatgagct gatataattt tcacaccctc atcactgaga tctctcccat caggaatggg 11640

tcagggagct cacaggtggc agcaactgct attacaggcc tcatctctac cagctcctgg 11700

ggcctgccct cctcccatta gaaaatcctc cacttgtcaa aaaggaagcc atttgctttg 11760

aactccaatt ccaccccсaa gaggctggga ccatcttact ggagtccttg atgctgtgtg 11820

acctgcagtg accactgccc catcattgct ggctgaggtg gttggggtcc atctggctat 11880

ctgggcagct gttctcttct ctcctttctc tcctgtttcc agacatgcag tatttccaga 11940

gagaaggggc cactctttgg caaagaacct gtctaacttg ctatctatgg caggaccttt 12000

gaagggttca caggaagcag cacaaattga tactattcca ccaagccatc agctccatct 12060

catccatgcc ctgtctctcc tttaggggtc cccttgccaa cagaatcaca gaggaccagc 12120

ctgaaagtgc agagacagca gctgaggcac agccaagagc tctggctgta ttaatgacct 12180

aagaagtcac cagaaagtca gaagggatga catgcagagg cccagcaatc tcagctaagt 12240

caactccacc agcctttcta gttgcccact gtgtgtacag caccctggta gggaccagag 12300

ccatgacagg gaataagact agactatgcc cttgaggagc tcacctctgt tcagggaaac 12360

aggcgtggaa acacaatggt ggtaaagagg aaagaggaca ataggattgc atgaagggga 12420

tggaaggtgc ccaggggagg aaatggttac atctgtgtga ggagtttggt gaggaaagac 12480

tctaagagaa ggctctgtct gtctgggttt ggaaggatgt gtaggagtct tctagggggc 12540

acaggcacac tccaggcata ggtaaagatc tgtaggtgtg gcttgttggg atgaatttca 12600

agtattttgg aatgaggaca gccatagaga caagggcaag agagaggcga tttaatagat 12660

tttatgccaa tggctccact tgagtttctg ataagaaccc agaacccttg gactccccag 12720

taacattgat tgagttgttt atgatacctc atagaatatg aactcaaagg aggtcagtga 12780

gtggtgtgtg tgtgattctt tgccaacttc caaggtggag aagcctcttc caactgcagg 12840

cagagcacag gtggccctgc tactggctgc agctccagcc ctgcctcctt ctctagcata 12900

taaacaatcc aacagcctca ctgaatcact gctgtgcagg gcaggaaagc tccatgcaca 12960

tagcccagca aagagcaaca cag 12983

<210> 2
<211> 9715
<212> DNA
<213> Homo sapiens

<400> 2

ctggttcatc tcattgggac tggttggaca agagggtgca gcccacggag ggtgagccaa 60

agcagggtgg ggcgtcgcct cacctgggaa gcacaagggg tcgtggaatt ttctcccta 120

cccaaggaaa gccataaggg actgagcctg aggaactgtg cactctggcc cagatactgc 180

acttttccca tggtctttgc aacccgcaga ccaggagatt ccctccggtg cctatgccac 240

cagggccctg ggtttcaagc acaaaactgg gcagccattt gggcagacac cgaactagct 300

gcaggagttt tttttttttt tttccatacc ccattggcac ctggaacgcc agtgagacag 360

aaccgttcac tcccctggaa agggggctga aaccagggat ccaagtggtc tggctcggtg 420

ggccccaccc ccatggagcc cagcaaacaa agattcactt ggcttgaaat tcttgctgcc 480

agcacagcag cagtctgaga ttgacctggg accctcgaac ttggttgggt gctgtggggg 540

ggcatcttcc attgctgagg cttgagtagg tggttttacc ttcgcggtgt aaacaaagct 600

gctgggaagt ttgaactggg tggagctcac cacagctcag taaggccact gtggccagac 660

tgcctctctg gatttctcct ctctgggaag gatatctctg aaaaaaaggc agcagcccca 720

gtcagggact tatagatgaa acccccatct ccctgggaca gagcccctcg gggaagaggt 780

ggcttccacc attgtggaag actgtgtggc aattcctcac ggatttagaa ctagagatac 840

catttgaccc agcaatccca ttactgggtg tatacccata ggattataaa tcattctact 900

ataaagacac atgcacactt atgtttattg taacactatt tacaatagca atgacctgga 960

accaatccaa aagcccatca atgatagact gaataaagaa aatgtggcac atatacactg 1020

tggaatacta tgcagccata aaaaaggatg agttcatgtc ctttgcagag acatggatga

1080

agctggaaac catcattctc agcaaactag cacaataaca gaaaaccaaa cactgcatgt
1140

tgtcactcat aagtgggagt taaacaatga gaacacatgg acacagggag gggaacgtca
1200

cacactgggg catgtcgggg agtgggggcc tacgggaggg atagcattag cagaaatacc
1260

taatgtaggt gacgggttga tgggtgcagc aaaccaccat ggcacatata cacctatgta
1320

ataaaactgc acgttctgca catgtacccc agaacttaaa gtataattaa taataataat
1380

aatttctggg catgtaagta gctgtctttc aggttctact ttgatacata ttctgagaga
1440

attaaacctg tcaaagaaac cttgactttc aatggcaggc actggaattg accctaataa
1500

tgtgttttgg ggtaagccta ctcatattct caacctgtct gcagtagtcg ttagaatctg
1560

aacttcctga agttcatgtg caaagttgag ttaattgttt aatattcaac aaggattatg
1620

ccagtaagat ggtaggaaaa tattagatat gtgtcatcac tgctggtatt atttaaactg
1680

caacatattt tagctggctg ctgatctcag ccaccatgcc tgcattttat ctctgtctcg
1740

tggtctgcaa ccttggaagc tttgaactta gctcatagaa tcctgggcat caagaacatg
1800

tggttctaat ggctagatag ggaatgagag taaaaggatt ttgcccacgg tcacgtgagt
1860

aaacaacaga tttggagggg tctggactac tgtgatgact tcattctgac aatatgttcc
1920

agttgtcctt tcatttcctc ctaatcacat gtctggtctg atctggctgt ttcccacctt
1980

ccaattcctg ccttctccaa tgctcccttc cgtaggtcac tctgtggctc agagaccctg
2040

cttagcaagc gcccaacctt tcaattattt gttcagtaaa acttgaactc atgtctcccc
2100

ttcttgataa aaagaaaata cgttatgtaa tgtcgggtta ctctataact cttgtcctgt 2160

ctctcggcaa ctactgaact aactgttttc atattgagca aacgtttatg gaaggactgc 2220

caagagtcag gtactaggct tggtaatatt ccccgttctc tctagtcaaa gccaacacca 2280

gccagacttg cagatctagg tcccaagccc actgcagatc acaggccagg gtctggtctc 2340

ctctgagctc ctttgggagg gaaagacaga attattaaca cccattttgt agattaggca 2400

actgaggctg aggaagttta aataactcag acagggcctg cacgtcagtc atattccaag 2460

gatccctact cactgtcttc tctctacaga acgagatgtc tctggagtcc atagaaagcc 2520

caggagcctg gctgggcacg gtggctcctg cctgtaatcc cagcactttg ggaggccgag 2580

gcaggcagat cacctgagct caggagttca agaccagcct gggcaacatg gcaaaacccc 2640

atctctacta aaaatacaaa aaattagctg ggcgtggtgg tgcatgcctc taatcccagc 2700

tacttgggag gctgaggcac aagaattgct tgagcccagg aggcagcagt tgcagtgagc 2760

tgagattgtg ccagtgcact ccagcctggg caacagagca agattccatt tcaaaaacaa 2820

aaacaaacac aaacaaacaa acaaaaatag aaagcccagg gaccacctgc gtcaggttcc 2880

cagccacacc tttttcttgt cctcctctgt ctctggcatc ttctcacagg ttcctaattg 2940

tttgtggttg cacaaattca aaatcccaga aaaattacca cttcacaccc actcagatgg 3000

ctattttttt tttgaaggaa gataacaagt gttgacaaga acatggagaa attggaattc 3060

tcacccattg ctggtgagaa tgtaatacgg tgctgctgct atggaaaaca gcttggagtt 3120

tcctcaaaaa gttcaacaga atttcaatgt gacccagcaa ttcccctcta agttatagat 3180

ctgagaggat taaaaacagt tactaaaata cacggactca catatttcta acagtccaat 3240

tcacaagggc caaaaggtgc taatagccca catgtccatc gatggatgga taaataaatt 3300

gtggtctatc catacaatgg aatattattc ggccataaat ggaatgaagt actgacgcat 3360

gctacagaat ggatgaaccg caaaaaaaat ggatgaacac atgctacaga atggatagcc 3420

tcactttact atgaagtgaa ggccagaaac gaagtccata tattgcatca tacaaaatat 3480

ccagaagagg gaagcccaca gagacagaat gtgcaatggt ggatgccagg gtctggggag 3540

aggggagagt ggggagaaac tgctcaactg gtacaggctt tattttggaa tgatgggaac 3600

attttgcaac tagatagagg tagtgattgc agaacacaga atgtactgaa ttccactgat 3660

ttttttcacc ttaaaatggt taattttcag tcctgagatt ggataatcat aaaaaaatgg 3720

ttaattttat gttatgtgaa tttcatccct atacatattt taaacctcag aaatatacac 3780

tagcaggcat ggaacaggtc actgtggtgc ctgccaagcc cggtgatgtt atctggggtc 3840

cccggccagc cttaagcctc ttgctgaccg gtggagggca gaacctttgc cctaaaagta 3900

taatatccac atgctggcat gattcctggc cagatggctt ctttattagc agtaattgaa 3960

actgcctcga tacagacact gtaccttgca accaaaaaat gactcaacaa tgataataag 4020

ggttaagctg ggcctttctc tctttgccag ttaaattata tttattatag cttgacatga 4080

aaaacaaagc aactccaaca ggtatcacaa gggcaaagga catgaacatt ttatcaaaga 4140

agaaatgcag ctgtcaaaaa tacagaaata ttcaaccttg ttcataataa agtggctggg

4200

ctcagtggtt catgcctgta atcccagtgc tttgcaaggc tgagacagga ggatcatttg
4260

aagccagaag ttcaagacca tcctaggcaa gtcagttcaa taccagactt catgtctaca
4320

aaacatcaaa aaattagcca ggcatggtga tgcatgcctg ttgtcccagc tactcaggag
4380

gctgaggcag gagaattgct tgagcctggg aggctgcggt ggcggtgagc catgattgtg
4440

ccattgtact ccagcctggg caatgcagca agactgtcta aataacaaaa ataatagtaa
4500

agaaaaggat tgggatgcca tttacttgcg tattcaatac acagagttaa aagtaatttc
4560

tacgttttct atttttttat tactaaaaaa agctggacca ttctcacagc ctgaaatgct
4620

tctcactttc ccttcttctg tccaaacact tctctatgat aatgcaaaca gtcactcctt
4680

taggaagact tcaccccagg tagttccaga tccccttatc tctgccttcc cagaactcct
4740

ggtgtctctc cagttccctc cgtgtggtga agtaccctac ctagggtttc agtatggctc
4800

tgtctgcaaa ggtcttgttc acaccttccc ttatggttct gttgccctgt gttgtgtcat
4860

agcacagggc acagtggaga acccattcac actgatagag agggccccat ggtcctggag
4920

ataaccatgt aaccgatcag aataaggcat tgagggctgg gtgtcaggcg tgggctgcac
4980

ttgggtgggc aggtcccctg gaaagtcact gggtttggca agcttcctag taacatgtct
5040

ctctggggtc cccctggaa cttcatgcaa aaatgctggt tgctggttta ttctagagag
5100

atggttcatt cctttcattt gattatcaaa gaaactcatg tcccaattaa aggtcataaa
5160

gcccagtttg taaactgaga tgatctcagc tgaatgaact tgctgaccct ctgctttcct
5220

ccagcctctc ggtgcccttg aaatcatgtc ggttcaagca gcctcatgag gcattacaaa 5280

gtttaattat ttcagtgatt attaaacctt gtcctgtgtt gaccccaggt gaatcacaag 5340

ctgaacttct gacaagaaca agctatcata ttcttttcaa ttacagaaaa aagtaagtta 5400

attgatagga tttttttgt ttaaaaaaa tgttactagt tttgaaaagg taatatgtgc 5460

acatggtaaa cactaagaag gtataagagc ataatgcttt tatactacta agaataatgt 5520

tttctctaag tttttttgg tagatgcttt catcagatta agaaaattcc ctgctattag 5580

ttgttgaagg tttttatatc ataaatgaaa gttgaatatt attatcatat attattaata 5640

tattgttatt gaactatcaa agccttttcc taaaaccatt gagatgatct tataaccatt 5700

ctcctttaac ctgttgacga gatcattggt atttatacta tttctctgtt aaccattctt 5760

gagtctcagg tttaaattca acttggtcat ggtgtgtcat ctttgatcat tgctgtctgt 5820

ggcttgctac tgttttgttt aggatttttg cactgatgct catcaatgag actggcatgc 5880

catcttcctt tgcagtcctg atttttttct gatttggatc atgtggttat ggccctcatg 5940

gaatgagttg ggcatgatgc cttttttca tgtctctgga ttgatgggac actttggatt 6000

ctctccagat ggccctcaat ggtccctgcc tcctcattgt taggcccctg ggcaagccct 6060

tctcatttct ggtaggccca ggaacctgtg ggggttttgt ttgtttgttt gtttcttgag 6120

tcggagtctc actctgtcac ccaggctgga gttggagtgc aatggcccga tcttggctca 6180

ctgcaacctc cacctcccag attcaagcaa ttctcctgcc tcagcctcct gagtagctgg 6240

aattacaggc acccaccgac acaccctgct aatttttgta tttttagtac agatggggtt 6300

tcacaatatt ggccaagctg gtctcgaact cctgatctca tgatctgccc ggcttggcct 6360

cccaaagtgt tgagattaca agcatgagcc accacaccca gtgaacctgt ggtttttaga 6420

agctccccat gcatgtgaat gctgtgagca tcccaggatg acagccactg tgtgttcagc 6480

tgttggaact gtgagaaagc accagtggga ccttctccag cacctgcctg ctgagttcat 6540

ggaagaggct tgttggggag atgatgccct ggctgactcc tgaaggatgg ttaggaatgc 6600

accagatgga agctgggttg gacccactct atgctgaaga acagcttgtg tggacacaag 6660

gagacacgga tatgtcattt ttgtagagcc tgaggagtgt ccaatcacac catttgctta 6720

aaacatcatg cacacttgga aaagtggact gagaccgaat gaagaagcta acagtggcca 6780

gatcagaaag ggtcttgtgt tacttcctag agatacttag attttatcct gtgggtgata 6840

ggagcagttg gagggactga agacaaggaa agaaacatgt ttcaagatct atgtttttca 6900

agacgctttt ctggtggctg agtagggaat tccctggata agtcctgccc agggtcaggc 6960

aaaacaagtt agggggttac tgaaataagg agtatgagaa atggtgtagg ttgtgctgac 7020

gttttgtaac acatctcatg atgatcttca tttccttcac taatttcctg tttcattaat 7080

tcccttccac gtgctcttct gaaatttgcc tcacattctc tgatttctct tttacctgtt 7140

ggtttcatca ccttttactt tttgctttcc tggaaacaca aatgattctg attgtgacat 7200

gtcagaatta tttgcaacat ttgcctttct gctgaaacca tgagttcact gaatacacaa 7260

tttagtaaag tgtaggatgc acatgtcgtt ttcgtggtca caaccagctc tgtagcattt

7320

tataactaca ctggcagtgt gctgggaggt gtagagagaa atatttatca catgtgtggc
7380

tgacacaacc tgccaagtta ttttaggagc ctccttggaa tcccagcaag aatgctaccg
7440

gcacaatttg taatcacagc atcctgctcc atgccttggc ttcatggcat agtcacttct
7500

gcaagtctct ttccagctgt ctgttcccat gtctataaag tatgagttaa atcatcctaa
7560

cactactcat cttacaaagt tttcttgctg atgttaagag agttgggaaa gaactgtata
7620

aactgtgaag tgccatggag atgttagtgg ttactttatc aagaaataga cactctagaa
7680

tggagtagaa agccaacagt tatgattgag tcctcctcct cttcttcttt ttattaattt
7740

ataaagaaaa gaggtttaat tgactcacag ttccatatgg ctggggaggc ctcgggaaac
7800

tctcagtcat agcaggaggc aaaggggaag aaggcacctt cttcacaagg cggcaggaga
7860

gagagagctc ctgttctttt ttgtcataaa gtctacagaa gtgcttatac ttcaggacaa
7920

gggcaggcag agagaaggaa ggacattgct tcaccccagc cctcactgac gagtttgcta
7980

ggggacctca ctttgtccca gagtagggca gaactctggc cactacccat tcagaaggcc
8040

tgggctgcac tgctagttcc tcactaactc tgtgtggcct tgggcaaggt tgggcctgtg
8100

ttaacagatt atgaccctgg gctctcaagc tagaggatct aaatttgaat cctggctctg
8160

ctaaagcaat tagtgatgta aactttaatg ggtcagttaa ccttcctgtg gcttagtttg
8220

ctcatctgta aaatagggat cataacagta tcaataccac atgattgttg gacagattga
8280

atcagttaat gcaggggaag tacttagcat gacacgtatt cactatcatt tcctggagta
8340

agagctgtgt gtgagtgggt gtgagcatgt gtgaaacctt ttctctgcaa tctcagttaa 8400

gaaaccaatc cagaatttaa agttcagggc ctaaatgggt ggttatcttc tcccagttcc 8460

atcctatccc acctttgctc ttcctcccgc ccacaggagc tgttggtcct tgattgggct 8520

ggaagacctg gtggacccta agtgatctat aagaggagaa tagagaacag ggaatgtctt 8580

caaaaatcta gagggacaca gaggctgaga ggcaggcagt cctgcagggt cttctgattg 8640

ggacaaggag aaccttggtc ttcacaggcc aattctggtc agtttccccc atggacagat 8700

gaggaaacag gcccaggaat atccaaggtc tcacacttcc catctgtcaa gtcttgttga 8760

ttctgttgta ttcatgtctc tcaaagggag atagagttta gggaagaaag aaggatcaac 8820

tgtgtctgat accactggga gcttaagtaa agggttcttt tacttcatag catttatccc 8880

aatttgtaat tcagtattat ttgtgtggct gtttggtgtc tctttctcct atatgagtgc 8940

tagcttcata agggcaagga ttttgattct ttaatattta gtgcttgcca catgccctga 9000

acacagcagg catacaggct aaccaacata cagtggcatg aaagtcatga aagtgagaca 9060

cctacctcct ccagtgccaa gagagcataa ccatgcacct gtcactctcc tcaacaccac 9120

ccccaagcat gaggcccaaa agcattagct aatcccctcc tccagccact aaaacttaaa 9180

ggccaggtgt ggtggctccc atctgaaatc ccagaacttc aggagacagc agcaggagga 9240

tcacttgagg ccaggagttt gagatcagcc tgggcaacat agctaggtcc catctgtact 9300

aaaaattagc tgggcgttgt tgcatgcctg tagtcccagc tactaaggag gctgaggtgg 9360

gaggatcact tgagcccagg aggtggaaac aacagtaagc tataatcaca gcactgaact 9420

ctagcctggg caacagagtg acaccctgcc tcaaaacaat tttaaaaata aataagagca 9480

aaacttagat accacgtggt caccccaaca tgcaaaatca agttttcccc tactgagaag 9540

aatggggact tgacagctga gttacagaga gataatcttc ttcttctttt tttttttttg 9600

gtttacatcc tcaagatcat gacttgtgaa atttgaatcg aatacacatg taattccaga 9660

gcaatgttgc ctccgcatac catcagcaat tcacttggct actggaagtc aggat 9715

<210> 3
<211> 632
<212> DNA
<213> Homo sapiens

<400> 3
tctagagaga tggttcattc ctttcatttg attatcaaag aaactcatgt cccaattaaa 60

ggtcataaag cccagtttgt aaactgagat gatctcagct gaatgaactt gctgaccctc 120

tgctttcctc cagcctctcg gtgcccttga aatcatgtcg gttcaagcag cctcatgagg 180

cattacaaag tttaattatt tcagtgatta ttaaaccttg tcctgtgttg accccaggtg 240

aatcacaagc tgaacttctg acaagaacaa gctatcatat tcttttcaat tacagaaaaa 300

agtaagttaa ttgataggat tttttttgtt taaaaaaaat gttactagtt tttgaaaagg 360

taatatgttg cacatggtaa acactaagaa ggtataagag cataatgctt ttatactact 420

aagaataatg ttttctctaa gttttttttg gtagatgctt tcatcagatt aagaaaattc 480

cctgctatta gttgttgaag gtttttatat cataaatgaa agttgaatat tattatcata 540

tattattaat atattgttat tgaactatca aagccttttc ctaaaaccat tgagatgatc 600

ttataaccat tctcctttaa cctgttgacg ag 632

<210> 4
<211> 11186
<212> DNA
<213> Homo sapiens

<400> 4
ggatccagtt tcagctttct acatatggct agccagtttt cccagcacca tttattaaat 60

agggaatcct ttccccattg cttgtttttg tcaggtttgt caaagatcag atggttgtag 120

atgtgtggtg tttgttctga ggcctctgtt ctgttccatt ggtccatatc cctgttttgg 180

tactagtacc atgctctttt ggttactgta gccttgtagt atagtttgaa gtcaggtagc 240

gtgattcctc cagctttgct ctttttgctt aggattgtct tgggaatgtg ggctcttttt 300

tggttccata tgaaatttaa agtagttttt tttccaattc tatgaagaaa gtcattggta 360

acttgatggg gatggcattg aatctataaa ttaccttggg aagtatggcc attttcacga 420

tattgattct tcctatccat gagcatggaa cattcttcca tttgtttgtg tcctctttga 480

ttttgttgag cagtggtttg tagttctcct tgaagaagtc cttcacctcc ctttaatttg 540

gattactaga tattttattc tcttagtaac aattgcaaat gggagttcac tcatgatttg 600

gctctctttc tgttattggt gtataggaat gcttgtgatt tttgcgcatt aattttgtat 660

cctgagactt tgctgaagtt gcttatcagc ttaaaaggat tttgggctga gacgatgggg 720

ttttctaaat atacaatcat ggcatctgca aacaggaaca atttgacttc ctcttttcct 780

aattgaatac cctttatttc tttttcttgc ctgattgccc tggccagaac ttccaatact 840

atgttgaata agagtcatga gtgagggcat cgttgtcttg tgctggtttc aaagtttttg 900

cccattcagt atgattttgg ctgtggtttt gccataaata gctcttatta ttttgagata 960

cgttccacca atacctactt tattgagagt ttttagcagg aagggctgtt gaattttgtc 1020

gaaggccttt tctacatcta ttgagacaat tatgtggttt tttaatcgtt gattctgttt 1080

atgtgatgga ttacatttat taatttgcat atgttgaacc agccttgcat cccagggatg 1140

aagcccactt gattgtagtg gataagcttt ttgatgtgct gctggattca gtttgccagt 1200

attttattga ggattttggc atcaatgttc atcagggata ttggtctaaa attctctttt 1260

tttgttgtgt ctctgccagg ctttggtatc aggatgatgc aggcctcaga aactgagtta 1320

gggaggattc cctcattttc tattgattgg aatagtttca gaaagaatgg taccagctac 1380

tctttgtacc tctggtagaa ttcagctgtg aatccatctg gtcctggact tttggttgg 1440

taggctatta attattgcct caattttagg gcctgttatt ggtctattca gacattcaac 1500

ttcttcccgg tttggtcttg ggagggttta tgtgtccagg aatttatcca tttcttctag 1560

attttctagt ttatttgtgt agaggtgttt atagtattgt ctgatggtag tttgtatttc 1620

tgtgagatcg gtggtgatat cccctttatc atttttatt gcatctattt aattcttctc 1680

tcttttcttc tttattattc tggctggcgg tctgtcaatt tttttgatct tttcaaaaaa 1740

ccagctcctg ggtttcactg attatttgaa gggtttttg tgtctctatt tctttcagtt 1800

ctcctgtgat cttagttatt tcttgccttc tgctagcttt tgaatgtgtt tgctcttcct

1860

tctctagttc tttgaattgt gatgttacag tgttgatttt agatctttcc tgctttctct
1920

tgtggtcatt tagtgctata aatttccctc tacacattgg tttacatgtg tctcagagat
1980

tctggtatgt tgtgtctttg ttctcattca tttcaagaac atctttactt ctgccttcat
2040

tttgttattt gcccagtagt cattcaggag caggttgttc agtcttcatg tagttgtgtg
2100

gttttgagtg agtttcttaa tcctgagttc taatttgatt gcactgttgt ctgagagaca
2160

gtttgttgtg atttccattc ttttacattt actgagcatg ctttatgtcc cattatgtgg
2220

tcaattttag aataagtgtg atgtgatgct gagaagaatg tatattctgt tgatttgggg
2280

tgtggagttc tgtagatgtc tattcagtcc actgggtgca gagctgagtg gacatgaaca
2340

ttttatcaaa gaagaaacac agctatcaaa aatccagaaa tattgaacct tgttaataat
2400

aaagtggctg gcctctggtt cattcctgta atctcagtcc tttgaaaggc tgagaaagga
2460

ggatcacttg aggccacaag ttcaagacca tcctagacaa gtcagttcaa gaccagactt
2520

catgtctaca aaacatcaaa aaattagcca ggcatggtga tgcatgcctg tcatcccagc
2580

tactcaggag gctgaggcag gaggattgct tgagcctggg agattgaagt ggcagtgagc
2640

catgattgtg ccattgcact ccagcctggg caatgcatca agactctgtc taaacaataa
2700

taataataat agtaatagta ataataataa taataaagaa aacggttggg acgccattcc
2760

ttacttattc aatacacaaa gttaaaagca atttctactt tctctatttt tttattacta
2820

aaaaaagctg aaccattctc acagcctgaa atgcttctca ccttcccctc ttctatacaa
2880

Untitled.ST25.txt

acacttctct gttgatgata atgcagacag tctctccttt aggaatactt cacaccaggt 2940

agttccagat ccccttatct ctgccttccc agagctcctg gtgtctcccc agttccctct 3000

gtgtggtgaa gtacccccac cttgggtctc agcatgactc gttctttgaa ggtcttgttc 3060

acattttccc ttatggttct gttcccctgt gttgtgtcac agcactgggc agagtggaca 3120

acccattcac accgatagag agggccccat ggttctggag ataaccatgt aactgatcag 3180

aatagggcat tgagggctgg gtgtcaggca tgggctgcac ttgggtgggc aggccccctg 3240

gaaagtcaca ggatttggca agcttcctag taacatctct ccctggggtc ctcttggaac 3300

ttcatgcccg atgctggatg ctggtttatt ctcgagagat ggttcattcc aataatcaat 3360

gaaactcatg tcccaactaa agttcataaa ctccagtttg taaactgaga taatctcagc 3420

tgaatgaact tgctgaccct ctgctttccc ccagcctctc agtgcccttg aaatcatgtc 3480

agttcaagca gccccatgag gcattacaat gtttagttat ttcagtgttt attaaacctt 3540

gccctatgct gaccccaggt gaatcacaag ctggacttct gacaaggaca agctatgata 3600

ttcttttcaa ttacagaaaa agtaagttaa ctgataggat tttttaaaga tgttactagt 3660

tttggaaagg taatttgtgc acatggtaaa caagaaggta taagaggata atgcttttat 3720

actgctgaga ataatgtttt ctctaatttt ttttggtaga tgctttcatc agattaataa 3780

aattcactgc tgttagttgt tgaaggtttt ttatatcatg aatgggagtt gaatattatc 3840

atgtattatt aatatattat tattgaacta gcaaaggctc ttcctaaaac aattgagatg 3900

atcttataat cgttctcctt taatctgttg atgagatcat tggtatttat actttttctc 3960

tgttaactat tcttgagtct caggtttaaa ttcaacttgg tcatggtgta tcatctttga 4020

acactcctgt ctctggcttg ctactattgt gttcagcatt tttgcactga tgccgatgaa 4080

tgagactggc atgtcatctt cctttgcggt cctgattttt ttcagatttg gatcatgtgg 4140

ccctcattga atgagttggg tgtgatgcct tctttttcat gtatctggat tgatgggaca 4200

ctttggagtc tctccagatg gccctcaatg gtccctgcct cctcattgtt aggctcctag 4260

gcaacccttt ctcatttctg gtaggcccag gaacctgtgg gttttatgtt tgtttgtttg 4320

tttgtttgtt tgtttttga gttggagtcc tgctttgtct cccaggctgg ggttggagtg 4380

caatggcctg atctcggccc actgcaacct ccacctcctg ggttcaagtg attctcctgc 4440

ctcagccttc tgtgtagctg ggattacagg catccaccac cactcctggc taatttttgt 4500

atttttagta gagacggggt tttacaatat aggccattgt gatctcttgg acaggctagt 4560

ctcaaattcc tgacctcatg atctgcctgc ctcagcctcc caaagtgctg agattacagt 4620

tttgtgcctc cacacacagt gaatctgtgg tttttaaaag ctcctcatgc atgtgaattc 4680

tgtgagcatc ccgggatgac agccactgtg tgtccagctg ttaaaactgt gagaaagcac 4740

cagcgggacc ctctccagca tttgcttgct gtggtcatga aagaggcttg ttggggagat 4800

gatgccctgg ttgactcctg aaggatggtt aggaatgcac cagatggaag ctgggttgga 4860

cccagtctat gctaaagaac agcttgtgtg gacacaagga gacacgaaca catcattttt 4920

gcagagcctg gggagtagcc aatcgcacca tttgcttaaa acaccgtgta cagttggaga

4980

agtggactga gacaggctga agaagctaac agtggccaga tgagaaaggg tcttgtgtta
5040

cttcctagat atacttagat tttatcctgt gagtgatagg aacagttgca gggactgaag
5100

ccaaggaagc atgctttaag attccatgtt ttttgagatg ctgtctggtg gctgagtagg
5160

gaattccctg gataagtact gcccagggta ggcaaaagaa gctaggaggt tactgaaata
5220

aggagtatga gaaatggtgt aggttttgct gatgttttgt aacacatctc atgacaatct
5280

tcatttcctt caccaatttc ctgtttcatt aattcccttc cacgtgctct tctgaaattt
5340

gcctcatatt ctttgatttc tcttttacat gttggtttca tcacctttta ctttttgctt
5400

tcctggaaac acaaatgatt ctgattgtga catgtcagaa ttatttgcaa cattcccctt
5460

tctgctgaaa catgagctca ctgaatacac aatttagtaa agtgtaggat gcacatgttg
5520

ttttcatggt cataaccagc tctgtagcat tttataacta cactggcagt gtgctgggag
5580

gtgtagagag aaatatttat ctcatgtgtg gctgacacaa cctgccaagt tgttttagga
5640

gccttcttgg aatcccagca agaacaccac tgatgcaatt tgaaatcaca atgtcctgct
5700

ccatgccctg gcttcatggc ttagtcacgt ctgaagtcta tttctaacta tctgtttcca
5760

catctataaa gtatgagtta aatcatccta atactactca tcttacaaag ttttcttgct
5820

gatattagga gagttgggaa agaactgtat aaattatgaa gtgccatgga gatgttggtg
5880

gttactttat caagaaatag acactccaga atagagtaga aagaaaacag ttatgattaa
5940

gtcctcctcc tcttcttttt ttttaattta caaagaaagg tttaattgag tcacagttcc
6000

atatggttgg ggaggctcag aaaacttgca atcatggcag ttggcaaagt ggaagaaggc 6060

accttcttca caaggtggca ggagagagag agctcctctt ctttttgtt gtaaagtcta 6120

cagaagtgca tatacttcag ggcaagggca ggcagggaga agaaaggaca ttgcttcacc 6180

ccagtcctca ctgacaagtt tgctttggga cttcattttg tcccagcata tgggacagag 6240

ctctggccac tacccattca gaaggcctga gctgcattgc tagttcccca ctaactctgt 6300

gtgtccttgg gcaaggctgg gcttatgtca aaagattatg accctgggct ctccagctac 6360

agaatctaca tatgaatcct ggctctgcta gagcaattag tgacgtaacc ttggatgggt 6420

cagttaacct tcctgtggct tagtttgctc atctgtaaaa tagggatcat aacaacatca 6480

ataccatggg ttgttagaca gattgaatca gttaatgcag ggtaaatact tagcatgaca 6540

cgtattcact atcatttcct tgagtaaaag ctgagtgtga gtgggtgtga gaatgtgtga 6600

aaccctttca ctgcaatctc agttaagaaa cccatccata atttaaagtt cagggcctaa 6660

atgggtggtt atcttctccc agttgcatcc tatcccacct ttgctcttct cctgcccgta 6720

ggagctgttg gtctttgatt gggctggaag acctggtgga ccctaagtga tctataagag 6780

aatgagaata gaggacaggg aatgtcttca aaactcctag agggacacag aggctgagag 6840

gcaggcagtc ctgcaggggt cttctgattg ggacaaggag gaccttggtc ttcataggcc 6900

aattctggtc aatttccccc atggacagat gaggaaacag atccaggaat atccaaggtc 6960

tcacacttcc catctgtcaa gtcttgttga ttctgttgta ttcatgtctt tcaaagagag 7020

agagagttta aggaaagaaa gaaggatcaa ctgtgtctga tatcactggg agcttaagta 7080

aagggttctt ttacttcata gcatttttcc caatttgtaa ttcagtatta tttttgtcac 7140

tgtttagtat ctctttgtcc tattagagag atagcttcat caggacaagg attttgattc 7200

tttaatattt agtgcttgcc acatgccctg aacacagcag gcatacagac taaccaacat 7260

acagtggcat cgaagtgaga cacctacctc ctccagtgcc tagagtacat gtccatggac 7320

ctgtcactct cctcaacacc acccctaagc atgaggcccg aaagcattgc taatcccctc 7380

ctccagccac caaaacttaa aggccaggtg tggtggctcc tatctgaaat ctcagaactt 7440

taggagacag cagcaggagg atcacttgag gccaggaatt tgagacgagc ctgggcaaca 7500

tagctagaca ccatctgtac taaaaattag ctgggcatgg tggtatacct gtagtaccag 7560

ctactaagga ggctgaggta ggaggatcac ttgaacccag gaggtggaag ctacagtgag 7620

ctataaccac agcactgaac tccagcctga gcaacagagt gagaccctgc ctcaaaacaa 7680

tttcaaaaat aaataaataa aaacaaaact tagataccac gtggtcaccc caacatgcaa 7740

aatcaagttt tcccctactg agaagaatgg ggacttgaga gctgagttac agagagataa 7800

tctgcctttt tttttttttt tttggtttac atcctcaaga tcatgacctg tgaaatttga 7860

atctaataca caaatcattc cagagcaatg ttgcttctgc ctaccacgag taattcactt 7920

ggccactgga agtcagaaca agcttcccag aagagaggta ccacttggac taccaatata 7980

aaaggatgaa aatatcggag tgaaggtgtt ccttgcatca ctgagtccct ggacagcctg 8040

tccactcatg ctgatatctg agcctaatgc ttctctgaat gttgagattt aactttgatc

Untitled.ST25.txt

8100

caatgaaacc agaccaagaa agaagaaacg tctttcattg ttgataagga catgattttt
8160

ctcacaattt tatgattatt tttccttagc tgtcctataa ttatctgctt atttgtctct
8220

tctccatgtg cttagggtac aaagttgacc aagaccaaga ataatgtctg ggagcacaat
8280

actgacagca cagctttaaa aacatgatga atgctttaat acaggaaatg agtaggggag
8340

aggcaagtgg tgcttgggtg ttcttccaat gcatagtatc ttccttgaca cagtcagtgc
8400

agctctcagt aggcaagtcc ctacatgtta gaagatgtta ctttctgtgg aattaggtgg
8460

cagaacatgc cttcaattat tttcctttgc agaacaacac caatttcatt agttaggaca
8520

gagtgctggc tgcatttgaa ttccaagcaa cgattagtct atcactgttg gtatagattc
8580

caaccagtca caccacctcc tgaagtttgt tgggcaggta aatcttcatc ttagaataaa
8640

aatcatctta gccaagtaag tgttttagag gaaagaagaa aacataatcg tttccataag
8700

agttttgttt ctaaaaaaat aagaaaggct ctttgtttag gtgagctaat gaagttgttg
8760

atagttatca gatgacactg gaatctttac ttgccagaat gtgttctgtg cacctctcgg
8820

tgtggcaaca tagagaggga gatcctccag caatgccatt gatatggtca gaaactgcat
8880

ctttctttct ccctgctgag atggggtcct ttgttctaga aaacccaggg ggtgccactg
8940

ggagtaaccc ttgagacagg aacacgaatc tcaaccaatt tctggttgca gccttgagtc
9000

ttactatttg ccatagtgat gcttagcaag gaatggcagg tgcaccagag cagcagagga
9060

cctaatatct cccttcctgt taacttttta taatatttta ttgtgatcag tatcagttgg
9120

gaagctactt gcagtcactg agcctcagtt tctacatctg taaactgggg atagtagcat 9180

ggccctattt aatgtgctca gcgaagccac tgaaaggaga cagaaatgta ccagaattcc 9240

ctggactttt atcctacttc tcctggggat tgtcacccac ctacccgtgt ctgtcctttg 9300

ttgctttgac gctgtcactt cttttcttag gtacctctct gtagggctcc attattccag 9360

ggattccaga gttacagcac atgcatacct ccatccaagc atgtttattt gtctcctgct 9420

tcactaggct gtccccaagg aacatgtggc tcccggcaca tacctggcac aacactgcac 9480

atgacattca cccacttggc cttgaatctg acaaggaatc tggcatgatg ttcacctgct 9540

gaggccaggt gccgagcagc cctggaggct taggggccag agggatggga aaaggtgtct 9600

ttctggggtg agtatcagtt tctgcaggag tgctgaacct gagaaagaat aaagagagaa 9660

ggaagtgaac aagcacagct taaacatcat ctgtttctac tgagttttaa caactctgag 9720

attttgtttg tcatggaatc catttctcag gccaagcaga cacagaactt gggtgtgagt 9780

gatgataatg agctgatata attttcacac cctcatcact gagatctctc ccatcaggaa 9840

tgggtcacag ggctcacagg tggcagcaac tgttattaca ggcctcatct ctaccagctc 9900

ctggcacctg ctctcctctc attagaaaat cctccacttg tcaaaaagga agccatttgc 9960

tttgaattcc aattccaccc tcaagaggct gggaccacct cattggagtc cttgatgctg 10020

tgtgacctgc agtgaccact gccccattgt tgctggctga ggtggtttgg gtcaacctgg 10080

ccatctgggc agctgttctc ttctcttctt tctcccctac tgtttccaga catgcagtat 10140

ttccagagag aaggggccac tctttggcaa agaacctgtc taactttcta tctacggcag 10200

gacttttgaa agctacagag gaagaagcac aaattgatgc tattccacta agccatcagc 10260

tccatctcat ccatgccatg tctctttttt aggggtcctc ttgccaacag aatcacagag 10320

gacaaatctg aaagtgcaga gacagcagct gaggcacagc caagagctct ggctgtatta 10380

atgacctaag aagatggagt ggtcaccaga aagtcagagg aagtgacaca caggggccca 10440

gcaatctcag ccaagtcaac tccaccagcc tttctggtcc ccactgtgtg tacagcaccc 10500

tgatagggac cagagccatg agagtgagta agaccagact atgcccttga ggagctcacc 10560

tctgctaagg gaaacaggcc tggaaacaca caatggtggt aaagaggaaa gaagacaata 10620

gaactgcatg aaggggatgg aaagtgccca ggggaggaaa tggttacttc tgtgtgaggg 10680

ggttggtgag gaaagactct aagagaaggc tctgtctggc tgggtatgaa aggatgtgta 10740

ggagtcttct agggggcaca ggcacactcc aggcataggt aaagatctgt aggcatggct 10800

tgttgggatg agtttcaagt attctggaat gaggacagcc atagagacaa gaggagagtt 10860

aatagatttt atgccaatgg ctccacttga gtttgtgata agaacccaga accccttggac 10920

tccccagtaa cattgattga gttgtgtatg attctacata gaatattaac tcaatggagg 10980

tcagtgagtg gtgtgtgtgt gattatttgc caactgccga ggtggagaag cctcttccga 11040

ctgcaggcag agcacggggg ccctgctact ggctgcagct ccagccctgc ctccttctcc 11100

agcatataaa caatccaaca gcctcactga atcactgctg tgcagggcag gaaagctcca 11160

cacacacagc ccagcaaaca gcagca

11186

<210> 5
<211> 10200
<212> DNA
<213> Homo sapiens

<400> 5
gccaattaga agaaacacaa ctacaaggtc agggcatatt attcaaacag tagagacaat 60

acagtcaaat atttggcaga attacaaaat atctcattgg aaaagacacg caagggaaat 120

caacaaaaag atatgaatca gaattcatct gtgtctcaag aaaaggtcat gcgataaatt 180

aagttctgct agtgtttcta cactaccgtt agcctcatta ccttattttt taagtgttaa 240

tatagtttta ggtattttac atacattttt attattaatt acaaccaaag tgcaacttgt 300

aatagcaatt ccttcacatt tttttttca aatcttgcac cttaaaatcc acctcgggcc 360

tcagttggcc agctttggta tctgatactt ggactacaga taccactaag gcaagtagat 420

aaaatgtact ctaggaccta cagcccttct gctagatcct gaagaatgat cattaaaaca 480

agctggtcta gctggtcaag agcaaaaata aaatcaagat gacagaaaat tgatgcaaaa 540

gtgaagtaaa atagctagag aatatgattg cgcctgtccc cttagcatgg attcccatgc 600

tagccaatct aaaatcctca ctgttagaat cctcctgtca atatgataga atgaacagca 660

agctcagtgt cagaaaacct gtgttgttaa cttggccctc tttctagctg aatgtgtgtt 720

tttggtcaag ttctttggca tttcagagac tcagagtagt gaaggaagtg gataagatga 780

cctctacatt ctcttgcaag ctcaaacatc tatgaatcca gagagaaaaa ctagagcatg 840

aaattaaggt tattttaaag aaataacctt aaaattatta gtattcgagg atctccaata

900

tattcatggc accactcaaa actttccttc tgctctatcc cgtcttggct caaagttatc
960

tccttaatga ggtctgccct gactatccta cttaaaattg taaactttgc ccacctggta
1020

cttccactct ctttcccctg ctctgttttt caccgtaata ctttactctt tttaacatac
1080

aaaatcactt atttactgtg ttgttatcta tctgcctact cttaccatca aatataagtt
1140

ctacctaggc agggattttt gtatgttttg ctcatggata tatacgaagc acttagagta
1200

atatgtgaca tatacagggt acttgattaa tactgttgag tgaatgaatg agtttccaat
1260

acaaatttaa aataaaatat ttcctaactt aaaattgtaa agtcagatct aaccaactgt
1320

tcattggtct gctagcagtg tttcttgtat atggaaatat attttaaata gatatgtcct
1380

gtgaaataat actaagtgtt ctaaagaaat aagtgagtga acgttacctc attgaactaa
1440

cttgaccttg ctcctgggag agagttcatt tgagattaaa caagttcaaa gtctatgaat
1500

cataaaacga taaaaaaaac taaaagggaa atggtgtttt tataagctct gcaattcaaa
1560

agccatttcg ggtaatattg ttatttttat gtcaggaatt cctcagtgct gatatcttag
1620

ggcaaagggt ttggttataa attaagagaa tgaggaaata ggtacatagt aggattgttc
1680

caaccaaata tgtgttgaat gtcaaaggaa tttccctgag gaataatctt cagaataatt
1740

tgctaagcac aggagaaaat ttggcttatt actttatagc cagatttcat ttttaattga
1800

aacttctttc aagcaaatca cttactagtc tattaacaat aacaacataa acacaagtaa
1860

acattcggaa tatagacatc caggtactaa gctgattgct ttacactcac tgtcttattt
1920

tacaagtaag gagttttagt tgcagcaaaa gaaataaatt ttccaatgtc aaatgaccag 1980

aacttaaacc caatctgttt ggtgctaaag ccaatgttct ttactgcaat gttgggttat 2040

cttgtttcta aaacttaaat ttatcagtaa aaggcaaaat ttgctattat tgaggacatt 2100

aaaatcatat ttttgtagac tctgaggaca aatccaacaa aaaagttcca actatttctt 2160

ggcaggcatc attgaaattg gtatatagct tccttgggta ttgactttga aaaggaagtt 2220

ggtcacttta gatatataag ttcagtctgt ttgtaaaaac aaaatgaaaa caaaacagtt 2280

gccttatatg ctaaaattat cctaatcgtt ttcaccttta acaacatata cacacagaac 2340

ttgaggaact ttacacggct catcttcata ttgtcagcat ctagcaaagt acttgccaca 2400

tagtgatcaa taaaagtttt agccagcctg ggcaacatag tgacagccta tctctacaaa 2460

aaaaaattag ccaggcaagg tggcgcacac ctttggtccc agctacttga gaggaggatg 2520

tggggagatc ctttgaatgc aggaggttga ggctgcagtg agctgtgatt gcgccactgt 2580

gcttcaacct tggtgacaga gcaagaccct gtctcacaca cacacaaatt aagtaaagaa 2640

aaaaaaaaga atcaaagaaa aaaataattc cccagcttaa gtccatcttt atttgtttgg 2700

ataagctata aagtgtcaaa taatgctgtt aatggacatt tctctagctc tcccaaagga 2760

ggaattgagc acatagtatg tgctgtattt tatatacaga ataaaaatag agacaagatt 2820

tctaccctca cagaacttaa attcttcagg agaatgacac tgaagtcctt aattggactt 2880

ctctcttctg tattatcttc ctcaagtgga ggtatatggt gcttagttat gaaaaatacc 2940

Untitled.ST25.txt

tccagggctt tgatcttctc aataactctt tgaggctgat atgaaaacag taattagaaa 3000

aaaaccatgt atccaactta tatagacagt tgatgaccaa agctagaatc cagttatttc 3060

agcctcccat gtattttctt attacttaag gagaatctct atctctacct ctttctctct 3120

tcttcctctc tcacttttct tagaaacatg ggtaagattt tcagaaatat gagaaactta 3180

ttaataaatg aaaaatactg ggaattctca atgtttcttg ttttagccag ttaattttgg 3240

ccttcattca atgtgagtgt cccttaataa ggagcaaact ccactgagag atagatacta 3300

ataaccagga ttctgaaaat gcattctcat ccccatctcc aaacttttat aaaaaatatt 3360

ataaaataat acacttttaa tataggaaat ttctcaaata cagaaaaaat taaagtaaac 3420

tcagacctaa ctttcatcac taaaagataa tcactcttga cactttgata tctttatctc 3480

taatattaca ccaattaatt tgcttgatat agtgaatatt atgctattat aattttcccc 3540

tgccttttgt ccttgcatat tagcataggt attttctgag gttatcacaa actctgtaag 3600

cacgttttat attactactt ttttaaagag gatgtataat aattaattca tccatatata 3660

tgtggttaag tattcaggtc actgctcatt tttcactgtt ataaaataaa gcagcaatga 3720

atacctttgg ctgatatttt tttctgtact tggaattatt tctttaagat agatttccta 3780

aaattgaatt actgagtcaa acagacttaa gttttcctta tgtatgtttc cttattcatt 3840

tgaataattt tcaactccta cttgtttatt taactcttgt gagcatgtga tagtctcatt 3900

tttcaaaata tctttgctgt tgtaatttgc atttctttgt agttagcatg aatatttcag 3960

tatgttttct tctgtgtacc agtatactac atactttttt atatgaattg cctatttgca

4020

tcttttgctc ggttctatta gacctttgaa ttttttctta tccatttata taagctcttt 4080

atatattaag aatattaacc tattgtgata tttgcaataa atagctatat ggtttgttgt 4140

tggttttaaa tgtgaattta ttcaattttc ttcataattt tgttgttttt atagatttct 4200

ttaaaagtaa taaaattatt gccatattat tatttatttt caatgtccat tactagccct 4260

ttccagtcat cacttttact ctcatgttct taatttttat tcatatcttg gctccattga 4320

ccatctttat taggatattt gggaaataat acaatttata ctaaacacac atcaaatctt 4380

accttatttt cttcttacaa aaaccctagg aatatgctgt ttttgtcttt atttgaatga 4440

cacagaaatc aaggtttttg agcagtggag tatttcttca aatgacacag aaatcaagtc 4500

ttttgagcag tggagtattt cttcaaagcc acacagctag taagtcatga agctggaatt 4560

ccaagagttg ccacttcatt ttcttctttc cctttatctt actcagttgt cttctctcct 4620

cttaattttg tcattcattt aaaaacattt cttgtgctat tatggtagat ttattttaat 4680

agggggcagt gacttactca gagagatgat tctctaatgg agttttaaag atcttagaag 4740

ttgatagagg aggctgggcg tggtggctca tgcctataat cccagcactt tgggaggctg 4800

aggtgggtgg atcacttgag gccagaagtt caagatcagc ctagtcaatg tggtgaaacc 4860

ccatttctac taaaattaca aaaattagcc tggtgtgctg gtgcactccc ataatcccag 4920

ctactcagga ggctgaggca tgagaattgc tggaacccag gaggcggagg ttgcagtgcg 4980

cctggattat taacactaca ctccagcctg ggagacagag taatactcca tctcaataaa 5040

aagaagttga taagggagat agttcatggc aacggatctt tgaaggcacg ctaatgataa 5100

cttaggcatt tagcctacta gtgtaatttc cataaatctg cctctgatgt catactctca 5160

gcacctaata ttttctacaa acatttattg aaactttatt ttgtataagt ctctgtccag 5220

tttgaatatt taaaaaattc ataatcatat gaaacattaa taataaatac aaaatgagag 5280

atgccgatac tgaaaagtag gattgcggag tggtagaaaa tatttctggc tgtagtagat 5340

ggggaagtgt tcaaagagga gtataattca ggtttccatt tgccatcgac ttatcacatg 5400

gctaactcac taagcgactt aattaaaatt aaattaattt atcatcatct gatcaccatt 5460

tcacacaact catgtctgtt gctgtattgg ctaaatgatg gcaagacaaa cgacctctga 5520

aaatgatcct attgaccttc ggaatctgga tttttttttc aatgcaggtg tccatagaag 5580

caatctgatg taatccaaca tgagttcaag cacagtcatt taatatcccc tatcaagtac 5640

agtcatttaa tatcccctat catcacatgt ccttcataca taaaaatcat tacatgtgaa 5700

agggtggaga gtgtgtggat cccttattat tgtgttattg taacacaata acaatattgt 5760

gttattattg ttaacacaag tgagtcatat gtcttgctct ttggactgag tggaaacttg 5820

tattctttct ctgcctcagg tcaattaact tcattgaggt gagttgcatt ccttctttaa 5880

gcatgttgaa ccttcaatct ggactcagat gggctaaata gaggagctag gaaaaataca 5940

gaaaataaat tattagagag atcagagaaa gatacataag atttacaata aaagaattat 6000

gagaaaaaca tccaaaagaa ttaaaaacca taggagaagg aaaaataggt gaacagcttt 6060

ttaatttcta taaatgtgtt gttaatactc ataataaagg actcagagct gggatatgag 6120

aataataggt caaacgtata tggatacata gatgtgacta catacatgag ttgcaaagaa 6180

tgctaaggag ggcaaaaaga gattgagaag agggcattat tactaatata tagcaatgtt 6240

gaatgtttag ggtgtttcag gcactgtaca aatcttttaa atacacaaat cacttaatct 6300

tgccataaca ttagaagata ctatctactc tttaccaaaa aggtaactgt gggatagcag 6360

agttaagtaa cccgttcaaa cctatgcata ataatcagca gagatggtcc ttatctaagt 6420

ttttctgcct ttgaagtcca aatagtttaa tgcagccagg tactaaagaa gaaaactttt 6480

gtaaattagt ttagtttaat gatttacatg tggaaaagca cagagtgaaa agcacatatc 6540

atctgagaag cccagtgagt ttggctgaaa tggagtgaac atgtacatgt tgagggtgag 6600

gaagatgatt agagagaagt gatgtgttgt ggttcttaaa agctaagagg aaactgttag 6660

atatgatata gtctgtggca gggagccatt gcaggttttc caagatggac tcataaggag 6720

aaaacacttc gtgggactgg aaaaggcagt gaagtggtgt gtcctatgat aatgtactac 6780

agtgtaggat agtggttaag agtacagtta tgagagaggg actactggtt agcaccttac 6840

ctgctgtgtg actgggcaaa taatgcaaac ctcagtgtct tttattgtaa tatgggagta 6900

acaaaaatag taactacttc ataggattct tgtaaagatt aaatgactta atttctttga 6960

agtgcttggc agttcctgat aaatgaccag tagttaataa atgttagttg ttattattat 7020

cattatatat tattactccc atagatacat atagaacaga ctgcagcaga gaggcaaatc 7080

tttaatgttg tcagagtata gacaagttgg tgaaatggct acatgagagc ggaggacaag

7140

aaggtgcaga ttgtggcagt cacttcaaat ggaaatatca ccgcttgaat gaaggtatat
7200

gagtgtcaac ttgcaagggg accaggtagg tttcatcaga aattaaggaa gcttaaggag
7260

aacagccaag ttcagcttga cagaagtggt ggtggcacaa atgcaagact ggtgtctttc
7320

aagaaaccaa ggactgttga aagtagcaag agctagtttg ttttaggtcc atcatgtttt
7380

atattcacac tttcatgtca gtggagcaaa gaaatggaat acaatataat agaatggtag
7440

aatcttattt ttaaaatctg tgttattctg atctttaact tacttatatc tttgatagag
7500

atctttacct gatgctcaag attgtagaaa tagtataatc aacataacag tatagcactg
7560

tatttatatc ctgcactgtt tagggagggt ttaaggccat tcaaaaggat acataaaata
7620

caacaagatt acataaatga aaggtgagat aaagcaacaa agcaaaacaa aagtgaaaac
7680

agagatcata ggcacaaata agattaaaaa cgcatgtaat gaagatgaaa gcttttacat
7740

ttaccccaga tggaccacag ggttgttgtt aagcctttaa acagtgaaca atgctgtaca
7800

cttgcatatg caattagaac atgtggaaaa aatagtggcc tgttagaagc ctaattaaca
7860

atttgtgaaa aaaaaaaaaa aaaaaaaaaa aagaggccga gctgtagctc acgcctgtaa
7920

tccctgcact ttgggaggcc gaggcgggcg gatcacgagg tcaggagatc aagaccatcc
7980

tggctaacac agtgaaaccc agtctctacg aaaaatacaa aaaattagcc gggcgtggtg
8040

gcgggagcct gtagtcccag ctacctggga ggctgaggca ggagaatggt gtgaacccgg
8100

gaggcggagc ttgcagtgag ccgagatcct gccactgcac tccagcctgg gcgacaaagc
8160

Untitled.ST25.txt

aagactccgt ctcaaaaaag aaaaaagaaa gaaaaacaaa agaaaacttc attgtattgt
8220

aaggccaaga acaaaatata tcaagataag gaaaatttgt agtcaagaat agaaaaaaat
8280

tatggctttg aagtatgagt tatttaaaga aagtggaaac atcctcagac tatgcagtaa
8340

aaaacaaagt gattttcttc ttctaaactt atgcaataaa ctgataggta atatgtgaaa
8400

gtcatagaat gtagactaga ggatacaaca aacctatttc ctctatgttc ataagaagta
8460

agaaaagctc tgatgtgagt tagcattgct ttacaatttt gaattgtgca gattgcacgt
8520

acttttcctc agtttgaagt aaatagtgga caggaaaaaa tattaaatgt tggcagtaaa
8580

tatggaagga aattacaact aatgtaatat gctaaaacat gctatgttta ttttactaat
8640

ttgaattaaa atgtaagaat ttaaaatgcc ctggaaaaac acgggcattg atctgacgtc
8700

tgaagtttta aaatattaca cactttgaaa tagcatttgt accttgaaat acctgtctct
8760

atatattttt taaaacttcc tttttctttc attccattta tcatcaaata aaggatgaac
8820

agatgtaact cagaaactgt caagcatgct gaagaaagac cactgcagaa aaatttctcc
8880

tagccttttc aaaggtgtta ggaagcagaa aggtgataca gaattggaga ggtcggagtt
8940

tttgtattaa ctgtattaaa tgcgaatccc gagaaaattt cccttaacta cgtcctgtag
9000

ttatatggat atgaagactt atgtgaactt tgaaagacgt gtctacataa gttgaaatgt
9060

ccccaatgat tcagctgatg cgcgtttctc tacttgccct ttctagagag gtgcaacgga
9120

agccagaaca ttcctcctgg aaattcaacc tgtttcgcag tttctcgagg aatcagcatt
9180

cagtcaatcc gggccgggag cagtcatctg tggtgaggct gattggctgg gcaggaacag 9240

cgccggggcg tgggctgagc acagccgctt cgctctcttt gccacaggaa gcctgagctc 9300

attcgagtag cggctcttcc aagctcaaag aagcagaggc cgctgttcgt ttcctttagg 9360

tctttccact aaagtcggag tatcttcttc caaaatttca cgtcttggtg gccgttccaa 9420

ggagcgcgag gtaggggcac gcaaagctgg gagctactat gggacagttc ccaagtgtca 9480

ggctttcaga tttcctgaac ttggtcttca cgggagaagg gcttcttgag gcgtggatag 9540

tgtgaagtcc tctggcaagt ccatggggac caagtggggt tagatctaga ctcaggagct 9600

cctggagcag cgcccaaacc gtagtggcac tggaccatgt tgcccggagc gcgcacagcc 9660

cgcgcggtgc ggggacctgc tctctgagcc cgcgggcggt gggtgggagg aagcatcgtc 9720

cgcggcgact ggaaccggga gggagaatcg cactggcggc gggcaaagtc cagaacgcgc 9780

tgccagaccc ccaactctgc cttcgtggag atgctggaga ccccgcgcac aggaaagccc 9840

ctgcagtgcc catcgcggcc agagcagctg gggcatcaac ggcgggcgct ccctcttact 9900

gctctctggc ttcgacgggg gactagaggt tagtctcacc tccagcgcgc ctgaggctca 9960

tgcatttggc taatgagctg cggtttctct tcaggtcgga atggatcttg aaggggaccg 10020

caatggagga gcaaagaaga agaacttttt taaactgaac aataaaaggt aactagcttg 10080

tttcattttc atagtttaca tagttgcgag atttgagtaa tttatttcta gcctccagct 10140

ctgaaataaa tgacatgttg ttgtttttaa ttatttttaa gaaacgcaag ctagcctttg 10200

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU01/01407

A.	**CLASSIFICATION OF SUBJECT MATTER**
Int. Cl. 7:	C12N 15/63, 5/10, C12Q 1/68, A01K 67/027, A61K 49/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
SEE ELECTRONIC DATABASE BELOW

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
SEE ELECTRONIC DATABASE BELOW

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Medline, WPIDS, CAPlus. Keywords: CYP3A4, silencer, enhancer, promoter, regulator, regulate, regulating, regulon.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 99/48915, A (GLAXO GROUP LIMITED) 30 September 1999. See page 17 lines 13-18.	1-20
Y	Eur J Drug Metab Pharmacokinet (1997) 22(4):311-3. Ogg MS, Gray TJB, Gibson GG. "Development of an in vitro reporter gene assay to assess xenobiotic induction of the human CYP3A4 gene." See whole document.	1-20
	Note: for the Y indications, WO 99/48915 and Eur J Drug Metab Pharmacokinet (1997) can be combined together.	

[X] Further documents are listed in the continuation of Box C [X] See patent family annex

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search 20 December 2001	Date of mailing of the international search report 24 DEC 2001
Name and mailing address of the ISA/AU AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA E-mail address: pct@ipaustralia.gov.au Facsimile No. (02) 6285 3929	Authorized officer J.H. CHAN Telephone No : (02) 6283 2340

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU01/01407

C (Continuation).	DOCUMENTS CONSIDERED TO BE RELEVANT	
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Cancer Chemother Pharmacol (1998) 42 Suppl:S50-3. Kamataki T, Yokoi T, Fujita K, Ando Y. "Preclinical approach for identifying drug interactions." See whole document.	
A	Chem Biol Interact (1997) 107(1-2):93-108. Olsen AK, Hansen KT, Friis C. "Pig hepatocytes as an in vitro model to study the regulation of human CYP3A4: prediction of drug-drug interactions with 17 alpha-ethynylestradiol." See whole document.	
A	WO 99/61622, A (THE UNIVERSITY OF SYDNEY) 2 December 1999. See whole document.	

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/AU01/01407

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report		Patent Family Member			
WO	9961622	AU	40232/99	EP	1082437
WO	9948915	AU	32116/99	EP	1066320

END OF ANNEX